VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV MATEMATIKY

FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF MATHEMATICS

STATISTICKÉ PŘEJÍMKY SROVNÁVÁNÍM

SAMPLING PROCEDURES FOR INSPECTION BY ATTRIBUTES

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE — PETR BARTUŠEK
AUTHOR

VEDOUCÍ PRÁCE — doc. RNDr. JAROSLAV MICHÁLEK, CSc.
SUPERVISOR

BRNO 2012

Vysoké učení technické v Brně, Fakulta strojního inženýrství

Ústav matematiky
Akademický rok: 2011/2012

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

student(ka): Petr Bartušek

který/která studuje v **bakalářském studijním programu**

obor: **Matematické inženýrství (3901R021)**

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma bakalářské práce:

Statistické přejímky srovnáváním

v anglickém jazyce:

Sampling procedures for inspection by attributes

Stručná charakteristika problematiky úkolu:

Při výběrové kontrole jakosti výrobků se používají statistické přejímky srovnáváním. Tyto přejímky se používají všude tam, kde lze aplikovat výběrovou kontrolu, tedy především u souborů tvořených velkým počtem výrobků hromadně nebo sériově vyráběných, dále u výrobků u nichž prověření jakosti vede k poškození nebo destrukci výrobku. Kontrola jakosti je prováděna pomocí přejímacího plánu na základě Českých technických norem. Tyto normy jsou odvozeny ze statistických vlastností náhodného výběru, při jejich konstrukci se používají základní diskrétní pravděpodobnostní rozdělení (hypergeometrické, binomické a Poissonovo).

Cíle bakalářské práce:

V práci popište hypergeometrické, binomické a Poissonovo rozdělení a zaměřte se na popis jejich asymptotických vlastností, zejména konvergence pravděpodobnostních funkcí. Rychlost těchto konvergencí demonstrujte pomocí simulačních programů. Získané výsledky potom použijte ke konstrukci přejímacích plánů pro statistické přejímky srovnáváním. Účinnost přejímacího plánu charakterizujte pomocí jeho operativní charakteristiky. Funkčnost přejímacího plánu pak ověřte pomocí simulací.

Seznam odborné literatury:

[1] Klůfa O.: Opravné statistické přejímky. Ekopress. ISBN: 978-80-86929-45-3. 2009

[2] Statistická přejímka srovnáváním. Československá státní norma ČSN 01 0254. Vydavatelství úřadu pro normalizaci a měření. Praha, 1983

[3] Česká technická norma ČSN ISO 2859-1 01 0261 Statistické přejímky srovnáváním - Část 1: přejímací plány AQL pro kontrolu každé dávky v sérii. Český normalizační institut, 2000.

Vedoucí bakalářské práce: doc. RNDr. Jaroslav Michálek, CSc.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2011/2012.

V Brně, dne 17.11.2011

L.S.

prof. RNDr. Josef Šlapal, CSc.	prof. RNDr. Miroslav Doupovec, CSc., dr. h. c.
Ředitel ústavu	Děkan fakulty

Abstrakt:

Tato bakalářská práce se zabývá popisem statistických přejímek srovnáváním. V práci jsou popsána tři diskrétní rozdělení pravděpodobností – hypergeometrické, binomické a Poissonovo, včetně konvergencí jejich pravděpodobnostních a distribučních funkcí. Hlavním cílem práce je pak odvození přejímacího plánu pro statistické přejímky srovnáváním. Práce je doplněna o tři vytvořené programy naprogramované v prostředí Matlab.

Abstract:

This bachelor´s thesis deals with a description of sampling procedures for inspection by attributes. In the thesis there are described three discrete probability distributions – hypergeometric distribution, binomial distribution and Poisson distribution, including the convergence of their probability mass functions and cumulative distribution functions. The main aim of the thesis is a determination of a sampling plan for sampling procedures for inspection by attributes. The thesis is supplemented with three created programs, which are programmed in software Matlab.

Klíčová slova:

Statistické přejímky, přejímací plán, rozdělení pravděpodobností

Keywords:

Sampling procedures, sampling plan, probability distribution

BARTUŠEK, P. *Statistické přejímky srovnáváním.* Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2012. 49 s. Vedoucí bakalářské práce doc. RNDr. Jaroslav Michálek, CSc.

Prohlašuji, že jsem bakalářskou práci *Statistické přejímky srovnáváním* vypracoval samostatně pod vedením doc. RNDr. Jaroslava Michálka, CSc. s použitím materiálů uvedených v seznamu zdrojů.

Petr Bartušek

Děkuji svému školiteli doc. RNDr. Jaroslavu Michálkovi, CSc. za rady, připomínky a čas věnovaný při konzultacích k této bakalářské práci.

Petr Bartušek

Osnova

1. Úvod

1.1 Cíl bakalářské práce

Cílem bakalářské práce je popis statistických přejímek srovnáváním a odvození konstrukce přejímacího plánu pro přejímky s jedním výběrem. Statistické přejímky se využívají při výběrové kontrole jakosti výrobků, jsou tedy velice důležité, pokud odběratel odebírá výrobky od dodavatele a potřebuje si ověřit jejich kvalitu. Podrobně jsou v této práci popsány statistické přejímky s jedním výběrem, při nichž dochází k rozhodnutí, o přijetí či zamítnutí dodávky, po jednom náhodném výběru.

Na začátku bakalářské práce je uvedeno označení a použitý matematický aparát, ze kterého se v dalších odstavcích vychází. Ve druhé kapitole jsou popsána rozdělení pravděpodobností, která jsou potřeba při odvozování konstrukce přejímacích plánů pro statistické přejímky srovnáváním. Jedná se o hypergeometrické, binomické a Poissonovo rozdělení. Závěr druhé kapitoly se zabývá vztahy mezi jednotlivými diskrétními rozděleními pravděpodobností a popisem vytvořeného simulačního programu v prostředí Matlab.

Třetí kapitola se týká statistických přejímek. Po vysvětlení základních pojmů a rozdělení přejímek, je největší část věnována popisu statistických přejímek srovnáváním s jedním výběrem. Pro tento typ přejímky jsou ukázány různé možnosti konstrukce přejímacího plánu a popis simulačního programu, pro výpočet přejímacího plánu (n, c), vytvořeného opět v programu Matlab. Přejímací plán (n, c) je dvojice čísel, podle nichž dojde po výběrové kontrole k rozhodnutí, zda přijmout nebo zamítnout dodávku výrobků. Na závěr třetí kapitoly jsou uvedeny další statistické přejímky srovnáváním, které se v praxi používají.

1.2 Označení a použitý matematický aparát

V tomto odstavci jsou uvedeny matematické pojmy a jejich označení (dle [1]), která se budou dále využívat.

Uvažujme pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$, kde Ω je prostor elementárních jevů, $\mathcal{A}$ je systém jevů definovaných na Ω (množinová σ-algebra) a P je pravděpodobnost.

Definice: Náhodnou veličinou X definovanou na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ rozumíme zobrazení $X: \Omega \to \mathbb{R}$ takové, že pro každé $x \in \mathbb{R}$ platí $\{\omega: X(\omega) \leq x\} \in \mathcal{A}$.

Definice: Nechť X je náhodná veličina na $(\Omega, \mathcal{A}, P)$. Pak funkci $F(x) = P(\{\omega: X(\omega) \leq x\})$, $x \in \mathbb{R}$, nazýváme distribuční funkcí náhodné veličiny X. Označení $F(x) = P(X \leq x)$.

Definice: Řekneme, že náhodná veličina X definována na $(\Omega, \mathcal{A}, P)$ je diskrétního typu, když existuje nejvýše spočetná množina $M \subset \mathbb{R}$ tak, že $P(X \in M) = 1$. Množinu M pak nazýváme oborem hodnot náhodné veličiny X a funkci $p(x) = P(X = x),\ x \in M,\ p(x) = 0,\ x \notin M$, nazýváme pravděpodobnostní funkcí diskrétní náhodné veličiny X.

Skutečnost, že X je náhodná veličina diskrétního typu s oborem hodnot M a pravděpodobnostní funkcí p, označme $X \sim (M, p)$.

Věta: Nechť X je náhodná veličina diskrétního typu s oborem hodnot M a pravděpodobnostní funkcí p, tj. $X \sim (M, p)$. Pak platí

$$F(x) = \sum_{\substack{t \leq x \\ t \in M}} p(t)\,.$$

Důkaz:

$$F(x) = P(X \leq x) = P([X \leq x] \cap [X \in M]) = P\left(\bigcup_{t \in M \cap (-\infty, x\rangle} [X = t] \right) =$$

$$= \sum_{t \in M \cap (-\infty, x\rangle} P(X = t) = \sum_{t \in M \cap (-\infty, x\rangle} p(t) = \sum_{\substack{t \leq x \\ t \in M}} p(t)$$

Definice: Nechť X je náhodná veličina, $F(x)$ její distribuční funkce a α je číslo z intervalu $(0, 1)$. Pak číslo x_α nazveme α-kvantilem rozdělení pravděpodobností náhodné veličiny X, když platí $x_\alpha = inf\{x: F(x) \geq \alpha\}$.

Pozn.: Při odvozování statistických přejímek se využívá zejména kvantil Pearsonova χ^2 rozdělení. Distribuční funkci Pearsonova χ^2 rozdělení o $2(c + 1)$ stupních volnosti označme $F_{\chi^2}(x; 2(c + 1))$. Jestliže platí

$$F_{\chi^2}\big(x; 2(c + 1)\big) = \alpha \,,$$

pak pro kvantil Pearsonova χ^2 rozdělení o $2(c + 1)$ stupních volnosti platí

$$x = \chi^2_\alpha[2(c + 1)] \,.$$

Definice: Střední hodnotou diskrétní náhodné veličiny $X \sim (M, p)$ rozumíme číslo EX, které je dáno výrazem

$$EX = \sum_{x \in M} xp(x) \,,$$

když uvedená řada absolutně konverguje.

Definice: Číslo $DX = E(X - EX)^2$ se nazývá rozptyl náhodné veličiny X za předpokladu, že uvedené střední hodnoty existují.

Definice: Nechť DX je rozptyl náhodné veličiny X. Pak číslo $\sigma(X) = \sqrt{DX}$ se nazývá směrodatná odchylka náhodné veličiny X.

Definice: Nechť X je náhodná veličina a existuje střední hodnota EX^k pro dané k, $k = 1, 2, ...$ Pak číslo EX^k se nazývá k-tým obecným momentem náhodné veličiny X a $\mu_k = E(X - EX)^k$ se nazývá k-tým centrálním momentem náhodné veličiny X.

Definice: Nechť X je náhodná veličina a existuje μ_3. Pak číslo

$$\alpha_3 = \frac{\mu_3}{\sigma^3(X)}$$

se nazývá šikmost rozdělení náhodné veličiny X.

Definice: Nechť X je náhodná veličina a existuje μ_4. Pak číslo

$$\alpha_4 = \frac{\mu_4}{\sigma^4(X)} - 3$$

se nazývá špičatost rozdělení náhodné veličiny X.

Definice: Nechť X je náhodná veličina na (Ω, $\mathscr{A}$, P). Pak funkci

$$M_X(t) = Ee^{tX}, \qquad t \in D_M \subset \mathbb{R},$$

nazýváme momentovou vytvořující funkcí náhodné veličiny X.

Věta: Nechť existuje k-tý obecný moment EX^k náhodné veličiny X. Pak existuje k-tá derivace $M_X^{(k)}(t)$ a platí

$$EX^k = \frac{d^k M_X}{dt^k}(0) .$$

Důkaz: $\mu_k = EX^k$ existuje pro k > 0. Důkaz se provede matematickou indukcí.
1) pro k = 1 platí

$$M_X'(t) = (Ee^{tX})' = E(e^{tX})' = E(Xe^{tX})$$

$$M_X'(0) = E(Xe^{0X}) = EX$$

2) pokud tvrzení platí pro k = m, pak platí pro k = m + 1

$$M_X^{(m+1)}(t) = (M_X^{(m)}(t))' = (E(X^m e^{tX}))' = E(X^{(m+1)} e^{tX})$$

$$M_X^{(m+1)}(0) = E(X^{m+1} e^{0X}) = EX^{m+1}$$

Dále pak podle [2] je zavedeno:

Definice: Norma vektoru je nezáporné číslo, které reprezentuje jeho velikost. Třída vektorových norem, známá jako l_p, závisí na parametru $1 \leq p \leq \infty$:

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} .$$

Pro p = 2 se l_2-norma nazývá Euklidovská norma a platí pro ni:

$$\|\boldsymbol{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} .$$

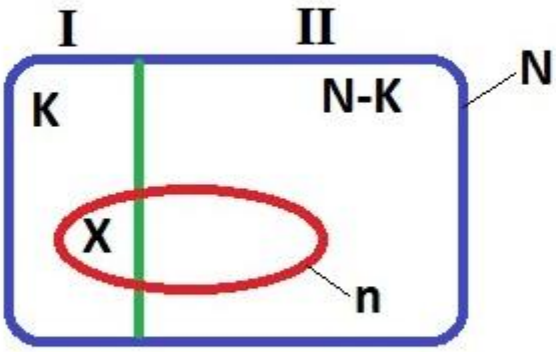
2. Rozdělení pravděpodobností používaná ke konstrukci přejímacích plánů

Při konstrukci přejímacích plánů jde o výběr z konečné populace. Při práci s ní a při konstrukci přejímacích plánů se používají diskrétní rozdělení pravděpodobností, konkrétně hypergeometrické, binomické a Poissonovo. V následujících odstavcích je u jednotlivých rozdělení uvedeno vždy označení, pravděpodobnostní funkce, číselné charakteristiky včetně jejich výpočtů a pro lepší představu jsou uvedeny grafy pravděpodobnostních funkcí pro různé hodnoty parametrů. Následně jsou ukázány vztahy mezi uvedenými diskrétními rozděleními pravděpodobností.

2.1 Hypergeometrické rozdělení

Urnové schéma

Celkový počet prvků je N, z nichž K je typu I a N – K je typu II $(\mathrm{K}, \mathrm{N} \in \mathbb{N}, 0 \leq \mathrm{K} \leq \mathrm{N})$. Z těchto N prvků je vybráno náhodně n prvků tak, že se vybrané nevracejí zpět (tzv. výběr bez opakování). Počet prvků typu I mezi n vybranými je náhodná veličina X. Obor hodnot náhodné veličiny X je $\mathrm{M} = \{\max\{0, \mathrm{n} + \mathrm{K} - \mathrm{N}\}, \max\{0, \mathrm{n} + \mathrm{K} - \mathrm{N}\} + 1, \dots, \min\{\mathrm{n}, \mathrm{K}\}\}$.
Grafické znázornění urnového schéma:



Obr.1 Urnové schéma

Pravděpodobnostní funkce náhodné veličiny X je určena vztahem:

$$p_{Hg}(x) = P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \qquad x \in M$$

$$p_{Hg}(x) = P(X = x) = 0, \qquad x \notin M$$

Říkáme, že náhodná veličina X má hypergeometrické rozdělení pravděpodobnosti s parametry N, K, n. Označení: $X \sim Hg(N, K, n)$.

Věta: Nechť náhodná veličina X má hypergeometrické rozdělení pravděpodobnosti, pak číselné charakteristiky jsou určeny vztahy:

- střední hodnota

$$EX = \frac{nK}{N}$$

- rozptyl

$$DX = \frac{nK}{N}\left(1 - \frac{K}{N}\right)\frac{N - n}{N - 1}$$

- směrodatná odchylka

$$\sigma(X) = \sqrt{\frac{nK}{N}\left(1 - \frac{K}{N}\right)\frac{N-n}{N-1}}$$

- 3. centrální moment

$$\mu_3 = \frac{nK(9N^2nK - 6nK^2N - 6n^2KN + 4K^2n^2 + 2N^2K^2 + 2N^2n^2 - 3N^3K - 3N^3n + N^4)}{N^3(N-1)(N-2)}$$

- 4. centrální moment

$$\begin{aligned}\mu_4 = {} & \frac{K(K-1)(K-2)(K-3)n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)} + \\ & + 6\left[\frac{K(K-1)(K-2)n(n-1)(n-2)}{N(N-1)(N-2)} + 3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}\right] - \\ & - 11\left[\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right] + 6\frac{nK}{N} - 4\left[\frac{K(K-1)(K-2)n(n-1)(n-2)}{N(N-1)(N-2)} + \right. \\ & \left. +3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}\right]\left(\frac{nK}{N}\right) + 6\left[\frac{K(K-1)n(n-1)}{N(N-1)} + \right. \\ & \left. + n\frac{K}{N}\right]\left(\frac{nK}{N}\right)^2 - 3\left(\frac{nK}{N}\right)^4\end{aligned}$$

- šikmost

$$\alpha_3 = \frac{(N-2K)(N-2n)\sqrt{N-1}}{(N-2)\sqrt{nK(N-K)(N-n)}}$$

- špičatost

$$\alpha_4 = \frac{(N-1)N^2[N(N+1) - 6K(N-K) - 6n(N-n)] + 6nK(N-K)(N-n)(5N-6)}{nK(N-K)(N-n)(N-2)(N-3)}$$

Důkaz: Je třeba stanovit číselné charakteristiky náhodné veličiny X, která má hypergeometrické rozdělení pravděpodobnosti. Při výpočtech se vychází z definičních předpisů. Obecné momenty EX^n jsou určeny podle vztahu

$$EX^n = \sum_{x \in M} x^n p(x) \,.$$

Přeznačme $p_{Hg}(x)$ na $p(x)$, pak postupně pro n = 1, 2, 3, 4 dostaneme:

$$EX = \sum_{x \in M} xp(x) = \sum_{x=0}^{n} x\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} = \sum_{x=1}^{n} x\frac{\frac{K!}{x!\,(K-x)!}\binom{N-K}{n-x}}{\frac{N!}{n!\,(N-n)!}} =$$

$$= K\frac{n}{N}\sum_{x=1}^{n}\frac{\frac{(K-1)!}{(x-1)!\,(K-x)!}\binom{N-K}{n-x}}{\frac{(N-1)!}{(n-1)!\,(N-n)!}} = \frac{nK}{N}\sum_{x=1}^{n}\frac{\binom{K-1}{x-1}\binom{N-K}{n-x}}{\binom{N-1}{n-1}} =$$

$$= \frac{nK}{N}\sum_{x=1}^{n}\frac{\binom{K-1}{x-1}\binom{N-1-(K-1)}{n-1-(x-1)}}{\binom{N-1}{n-1}} = \frac{nK}{N}$$

$$EX^2 = \sum_{x\in M} x^2p(x) = \sum_{x\in M}(x(x-1)+x)p(x) = \sum_{x=1}^{n} x(x-1)p(x) + \sum_{x=1}^{n} xp(x) =$$

$$= \sum_{x=1}^{n} x(x-1)\frac{\frac{K!}{x!\,(K-x)!}\binom{N-K}{n-x}}{\frac{N!}{n!\,(N-n)!}} + EX =$$

$$= K(K-1)\frac{n(n-1)}{N(N-1)}\sum_{x=1}^{n}\frac{\frac{(K-2)!}{(x-2)!\,(K-x)!}\binom{N-K}{n-x}}{\frac{(N-2)!}{(n-2)!\,(N-n)!}} + EX =$$

$$= K(K-1)\frac{n(n-1)}{N(N-1)}\sum_{x=1}^{n}\frac{\binom{K-2}{x-2}\binom{N-K}{n-x}}{\binom{N-2}{n-2}} + EX = \frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}$$

$$DX = EX^2 - E^2X = \frac{(K^2-K)(n^2-n)+Kn(N-1)}{N(N-1)} - \left(n\frac{K}{N}\right)^2 =$$

$$= \frac{-NK^2n - NKn^2 + N^2Kn + K^2n^2}{N^2(N-1)} = n\frac{K}{N}\frac{-NK - Nn + N^2 + Kn}{N(N-1)} =$$

$$= n\frac{K}{N}\frac{K(n-N)+N(N-n)}{N(N-1)} = n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}$$

$$\sigma(X) = \sqrt{DX} = \sqrt{n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}}$$

Pro výpočet 3. a 4. centrálního momentu je potřeba nejdříve určit obecné momenty EX^3 a EX^4:

$$EX^3 = \sum_{x\in M} x^3p(x) = \sum_{x\in M}[x(x-1)(x-2)+3x^2-2x]p(x) =$$

$$= \sum_{x\in M} x(x-1)(x-2)p(x) + 3\sum_{x\in M} x^2p(x) - 2\sum_{x\in M} xp(x) =$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} + 3EX^2 - 2EX =$$

$$= \sum_{x=1}^{n} x(x-1)(x-2)\frac{\frac{K!}{x!\,(K-x)!}\binom{N-K}{n-x}}{\frac{N!}{n!\,(N-n)!}} + 3EX^2 - 2EX =$$

$$= K(K-1)(K-2)\frac{n(n-1)(n-2)}{N(N-1)(N-2)}\sum_{x=1}^{n}\frac{\frac{(K-3)!}{(x-3)!\,(K-x)!}\binom{N-K}{n-x}}{\frac{(N-3)!}{(n-3)!\,(N-n)!}} + 3EX^2 - 2EX =$$

$$= K(K-1)(K-2)\frac{n(n-1)(n-2)}{N(N-1)(N-2)}\sum_{x=1}^{n}\frac{\binom{K-3}{x-3}\binom{N-3-(K-3)}{n-3-(x-3)}}{\binom{N-3}{n-3}} + 3EX^2 - 2EX =$$

$$= K(K-1)(K-2)\frac{n(n-1)(n-2)}{N(N-1)(N-2)} + 3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}$$

$$EX^4 = \sum_{x\in M} x^4 p(x) = \sum_{x\in M}[x(x-1)(x-2)(x-3) + 6x^3 - 11x^2 + 6x]p(x) =$$

$$= \sum_{x\in M} x(x-1)(x-2)(x-3)p(x) + 6\sum_{x\in M} x^3 p(x) - 11\sum_{x\in M} x^2 p(x) + 6\sum_{x\in M} xp(x) =$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)(x-3)\frac{\frac{K!}{x!\,(K-x)!}\binom{N-K}{n-x}}{\frac{N!}{n!\,(N-n)!}} + 6EX^3 - 11EX^2 + 6EX =$$

$$= \frac{K(K-1)(K-2)(K-3)n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)}\sum_{x=1}^{n}\frac{\frac{(K-4)!}{(x-4)!\,(K-x)!}\binom{N-K}{n-x}}{\frac{(N-4)!}{(n-4)!\,(N-n)!}} +$$

$$+ 6EX^3 - 11EX^2 + 6E$$

$$= \frac{K(K-1)(K-2)(K-3)n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)}\sum_{x=1}^{n}\frac{\binom{K-4}{x-4}\binom{N-4-(K-4)}{n-4-(x-4)}}{\binom{N-4}{n-4}} +$$

$$+ 6EX^3 - 11EX^2 + 6EX =$$

$$= \frac{K(K-1)(K-2)(K-3)n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)} +$$

$$+ 6\left[\frac{K(K-1)(K-2)n(n-1)(n-2)}{N(N-1)(N-2)} + 3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}\right] -$$

$$- 11\left[\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right] + 6\frac{nK}{N}$$

$$\mu_3 = E(X-EX)^3 = E(X^3 - 3X^2EX + 3XE^2X - E^3X) =$$

$$= EX^3 - 3EX^2EX + 3EXE^2X - E^3X = EX^3 - 3EX^2EX + 2E^3X =$$

$$= \left[K(K-1)(K-2)\frac{n(n-1)(n-2)}{N(N-1)(N-2)} + 3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}\right] -$$

$$-3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right)\left(n\frac{K}{N}\right) + 2\left(n\frac{K}{N}\right)^3 =$$

$$= \frac{nK(9N^2nK - 6nK^2N - 6n^2KN + 4K^2n^2 + 2N^2K^2 + 2N^2n^2 - 3N^3K - 3N^3n + N^4)}{N^3(N-1)(N-2)}$$

$$\mu_4 = E(X-EX)^4 = E(X^4 - 4X^3EX + 6X^2E^2X - 4XE^3X + E^4X) = EX^4 - 4EX^3EX +$$

$$+ 6EX^2E^2X - 4E^4X + E^4X = EX^4 - 4EX^3EX + 6EX^2E^2X - 3E^4X =$$

$$= \frac{K(K-1)(K-2)(K-3)n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)} +$$

$$+ 6\left[\frac{K(K-1)(K-2)n(n-1)(n-2)}{N(N-1)(N-2)} + 3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}\right] -$$

$$- 11\left[\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right] + 6\frac{nK}{N} - 4\left[\frac{K(K-1)(K-2)n(n-1)(n-2)}{N(N-1)(N-2)} +\right.$$

$$\left. +3\left(\frac{K(K-1)n(n-1)}{N(N-1)} + n\frac{K}{N}\right) - 2n\frac{K}{N}\right]\left(\frac{nK}{N}\right) + 6\left[\frac{K(K-1)n(n-1)}{N(N-1)} +\right.$$

$$\left. + n\frac{K}{N}\right]\left(\frac{nK}{N}\right)^2 - 3\left(\frac{nK}{N}\right)^4$$

$$\alpha_3 = \frac{\mu_3}{\sigma(X)^3} =$$

$$= \frac{\dfrac{nK(9N^2nK - 6nK^2N - 6n^2KN + 4K^2n^2 + 2N^2K^2 + 2N^2n^2 - 3N^3K - 3N^3n + N^4)}{N^3(N-1)(N-2)}}{\sqrt{n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}}^{\,3}} =$$

$$= \frac{(N-2K)(N-2n)\sqrt{N-1}}{(N-2)\sqrt{nK(N-K)(N-n)}}$$

$$\alpha_4 = \frac{\mu_4}{\sigma(X)^4} - 3 = \frac{\mu_4}{\sqrt{n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}}^{\,4}} - 3 =$$

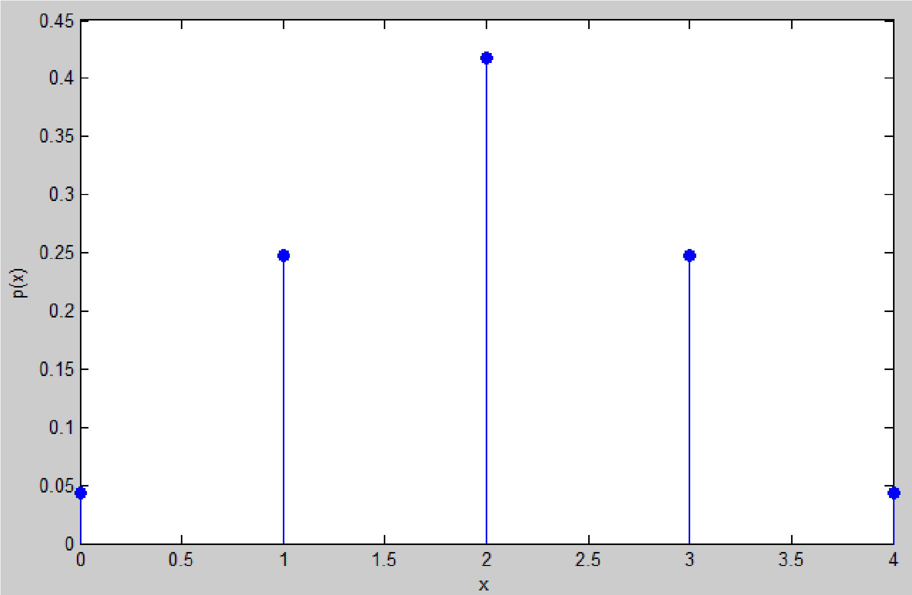
$$= \frac{(N-1)N^2[N(N+1) - 6K(N-K) - 6n(N-n)] + 6nK(N-K)(N-n)(5N-6)}{nK(N-K)(N-n)(N-2)(N-3)}$$

Pozn.: Důkaz lze provést i jinak, např. pomocí faktoriálních momentů.

Tvar pravděpodobnostní funkce pro hypergeometrické rozdělení s různými parametry je dále graficky znázorněn na obr. 2 a 3. Pro oba uváděné příklady jsou vypočítány číselné charakteristiky, které jsou následně porovnány.

Z obrázků a vypočtených číselných charakteristik je vidět, že pro $X \sim Hg(500, 35, 80)$ oproti $X \sim Hg(20, 4, 10)$ vzrostla střední hodnota (v grafu je posunuta více doprava), zvětšil se také rozptyl a směrodatná odchylka (širší graf). Pro $X \sim Hg(20, 4, 10)$ vyšla šikmost 0, což potvrzuje obrázek, ve kterém je pravděpodobnostní funkce symetrická kolem střední hodnoty. U obr. 3 vyšla šikmost kladná, což je znázorněno strmějším nárůstem a pozvolnějším poklesem hodnot. U obr. 2 se naopak projevuje výraznější špičatost, která má zápornou hodnotu.

Grafické znázornění pravděpodobnostní funkce s vypočtenými číselnými charakteristikami:



$EX = 2$

$DX = 0{,}8421$

$\sigma(X) = 0{,}9177$

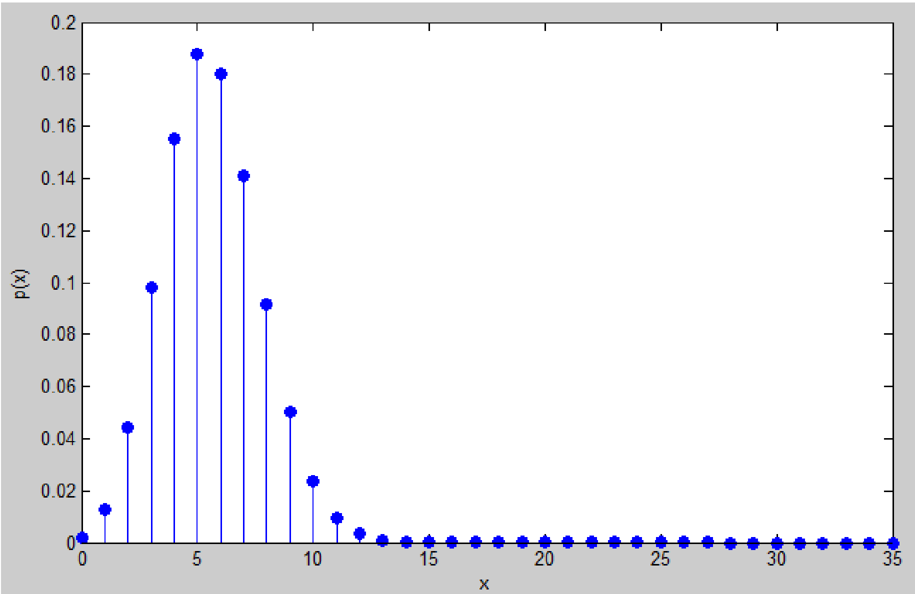
$\mu_3 = 0$

$\mu_4 = 1{,}8824$

$\alpha_3 = 0$

$\alpha_4 = -0{,}3456$

Obr. 2 Pravděpodobnostní funkce náhodné veličiny $X \sim Hg(20, 4, 10)$



$EX = 5{,}6$

$DX = 4{,}3835$

$\sigma(X) = 2{,}0937$

$\mu_3 = 2{,}5738$

$\mu_4 = 57{,}9432$

$\alpha_3 = 0{,}2804$

$\alpha_4 = 0{,}0155$

Obr. 3 Pravděpodobnostní funkce náhodné veličiny $X \sim Hg(500, 35, 80)$

2.2 Binomické rozdělení

Při počítání hodnot pravděpodobnostní funkce hypergeometrického rozdělení je potřeba pracovat s kombinačními čísly, a proto je někdy nutné vyčíslovat faktoriály z velkých čísel. Z důvodů výpočetní náročnosti je možné, při splnění určitých podmínek uvedených v kapitole 2.4, nahradit hypergeometrické rozdělení binomickým. Cílem tohoto odstavce je charakterizovat binomické rozdělení a popsat jeho charakteristiky.

Uvažujme alternativní náhodný pokus, který má dva možné výsledky – jev A (úspěch) a jev opačný $\overline{\mathrm{A}}$ (neúspěch). Pravděpodobnost úspěchu je P(A) = θ. Pokus n krát nezávisle opakujeme. Pravděpodobnost úspěchu θ je stejná pro všechny pokusy. Náhodná veličina X označuje počet nastoupení jevu A v posloupnosti nezávislých pokusů délky n. Náhodná veličina X má obor hodnot $\mathrm{M} = \{0,1,2, \dots, \mathrm{n}\}$.

Pravděpodobnostní funkce náhodné veličiny X je určena vztahem:

$$p_{Bi}(x) = P(X = x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \qquad x \in M$$

$$p_{Bi}(x) = P(X = x) = 0, \qquad x \notin M$$

Pak říkáme, že náhodná veličina X má binomické rozdělení s parametry n a θ: n ∈ ℕ, 0 < θ < 1. Označení: $X \sim Bi(n, \theta)$.

Věta: Nechť náhodná veličina X má binomické rozdělení pravděpodobnosti, pak číselné charakteristiky jsou určeny vztahy:

- střední hodnota $EX = n\theta$
- rozptyl $DX = n\theta(1-\theta)$
- směrodatná odchylka $\sigma(X) = \sqrt{n\theta(1-\theta)}$
- 3. centrální moment $\mu_3 = n\theta - 3n\theta^2 + 2n\theta^3$
- 4. centrální moment $\mu_4 = 3n^2\theta^2 - 6n^2\theta^3 + 3n^2\theta^4 + n\theta - 7n\theta^2 + 12n\theta^3 - 6n\theta^4$
- šikmost $\alpha_3 = \frac{1-2\theta}{\sqrt{n\theta(1-\theta)}}$
- špičatost $\alpha_4 = \frac{1-6\theta(1-\theta)}{n\theta(1-\theta)}$

Důkaz: Přeznačme $p_{Bi}(x)$ na $p(x)$. Pro výpočet hodnot EX^n je možno využít dvou různých způsobů výpočtu. V následujících výpočtech jsou ukázány oba možné přístupy. Hodnoty EX a EX^2 jsou určeny podle vztahu

$$EX^n = \sum_{x \in M} x^n p(x),$$

hodnoty EX^3 a EX^4 pak pomocí momentové vytvořující funkce.

$$EX = \sum_{x \in M} xp(x) = \sum_{x=0}^{n} x \binom{n}{x} \theta^x (1-\theta)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} \theta^x (1-\theta)^{n-x} =$$

$$= \sum_{x=1}^{n} x \frac{n!}{x!\,(n-x)!} \theta^x (1-\theta)^{n-x} = n\theta \sum_{x=1}^{n} \frac{(n-1)!}{(n-x)!\,(x-1)!} \theta^{x-1} (1-\theta)^{n-x} =$$

$$= n\theta \sum_{x=1}^{n} \binom{n-1}{x-1} \theta^{x-1} (1-\theta)^{n-1-(x-1)} = n\theta$$

$$EX^2 = \sum_{x \in M} x^2 p(x) = \sum_{x=0}^{n} x^2 \binom{n}{x} \theta^x (1-\theta)^{n-x} = \sum_{x=0}^{n} (x(x-1)+x) \binom{n}{x} \theta^x (1-\theta)^{n-x} =$$

$$= \sum_{x=1}^{n} x(x-1) \frac{n!}{x!\,(n-x)!} \theta^x (1-\theta)^{n-x} + \sum_{x=0}^{n} x \binom{n}{x} \theta^x (1-\theta)^{n-x} =$$

$$= n(n-1)\theta^2 \sum_{x=1}^{n} \frac{(n-2)!}{(x-2)!\,(n-x)!} \theta^{x-2} (1-\theta)^{n-2-(x-2)} + EX =$$

$$= n(n-1)\theta^2 + n\theta = n^2\theta^2 - n\theta^2 + n\theta$$

$$DX = EX^2 - E^2X = n^2\theta^2 - n\theta^2 + n\theta - n^2\theta^2 = n\theta(1-\theta)$$

$$\sigma(X) = \sqrt{DX} = \sqrt{n\theta(1-\theta)}$$

Pro výpočet 3. a 4. centrálního momentu je potřeba nejdříve určit EX^3 a EX^4. Zde se využívá momentové vytvořující funkce, která má pro binomické rozdělení podle [13] tvar

$$M_X(t) = (1 - \theta + \theta e^t)^n.$$

$$EX^3 = \frac{d^3 M_X}{dt^3}(0) = n^3\theta^3 + 3n^2\theta^2 - 3n^2\theta^3 + n\theta - 3n\theta^2 + 2n\theta^3$$

$$EX^4 = \frac{d^4 M_X}{dt^4}(0) = n^4\theta^4 + 6n^3\theta^3 - 6n^3\theta^4 + 7n^2\theta^2 - 18n^2\theta^3 + 11n^2\theta^4 + n\theta -$$
$$-7n\theta^2 + 12n\theta^3 - 6n\theta^4$$

$$\mu_3 = E(X-EX)^3 = EX^3 - 3EX^2EX + 2E^3X = n^3\theta^3 + 3n^2\theta^2 - 3n^2\theta^3 + n\theta - 3n\theta^2 +$$
$$+2n\theta^3 - 3(n^2\theta^2 - n\theta^2 + n\theta)(n\theta) + 2(n\theta)^3 = n\theta - 3n\theta^2 + 2n\theta^3$$

$$\mu_4 = E(X-EX)^4 = EX^4 - 4EX^3EX + 6EX^2E^2X - 3E^4X = n^4\theta^4 + 6n^3\theta^3 - 6n^3\theta^4 +$$
$$+7n^2\theta^2 - 18n^2\theta^3 + 11n^2\theta^4 + n\theta - 7n\theta^2 + 12n\theta^3 - 6n\theta^4 - 4(n^3\theta^3 + 3n^2\theta^2 -$$
$$-3n^2\theta^3 + n\theta - 3n\theta^2 + 2n\theta^3)(n\theta) + 6(n^2\theta^2 - n\theta^2 + n\theta)(n\theta)^2 - 3(n\theta)^4 =$$
$$= 3n^2\theta^2 - 6n^2\theta^3 + 3n^2\theta^4 + n\theta - 7n\theta^2 + 12n\theta^3 - 6n\theta^4$$

$$\alpha_3 = \frac{\mu_3}{\sigma(X)^3} = \frac{n\theta - 3n\theta^2 + 2n\theta^3}{\sqrt{n\theta(1-\theta)}^3} = \frac{n\theta - n\theta^2 - 2n\theta^2 + 2n\theta^3}{\sqrt{n\theta(1-\theta)}^3} = \frac{(1-2\theta)(n\theta - n\theta^2)}{\sqrt{n\theta(1-\theta)}^3} =$$

$$= \frac{(1-2\theta)n\theta(1-\theta)}{\sqrt{n\theta(1-\theta)}\,n\theta(1-\theta)} = \frac{1-2\theta}{\sqrt{n\theta(1-\theta)}}$$

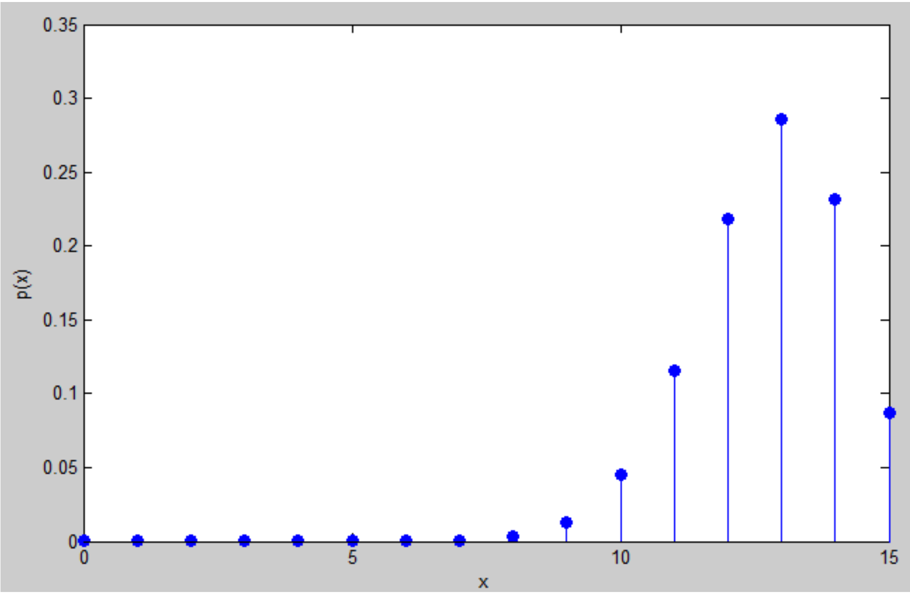
$$\alpha_4 = \frac{\mu_4}{\sigma(X)^4} - 3 = \frac{3n^2\theta^2 - 6n^2\theta^3 + 3n^2\theta^4 + n\theta - 7n\theta^2 + 12n\theta^3 - 6n\theta^4}{\sqrt{n\theta(1-\theta)}^4} - 3 =$$

$$= \frac{n\theta - n\theta^2 - 6n\theta^2 + 6n\theta^3 + 6n\theta^3 - 6n\theta^4}{n\theta(1-\theta)n\theta(1-\theta)} = \frac{[1 - 6\theta + 6\theta^2][n\theta - n\theta^2]}{n\theta(1-\theta)n\theta(1-\theta)} =$$

$$= \frac{[1 - 6\theta(1-\theta)][n\theta(1-\theta)]}{n\theta(1-\theta)n\theta(1-\theta)} = \frac{1 - 6\theta(1-\theta)}{n\theta(1-\theta)}$$

Tvar pravděpodobnostní funkce pro binomické rozdělení s různými parametry je dále graficky znázorněn na obr. 4 a 5. Pro oba uváděné příklady jsou vypočítány číselné charakteristiky, které jsou následně porovnány.

Grafické znázornění pravděpodobnostní funkce s vypočtenými číselnými charakteristikami:



$EX = 12{,}75$

$DX = 1{,}9125$

$\sigma(X) = 1{,}3829$

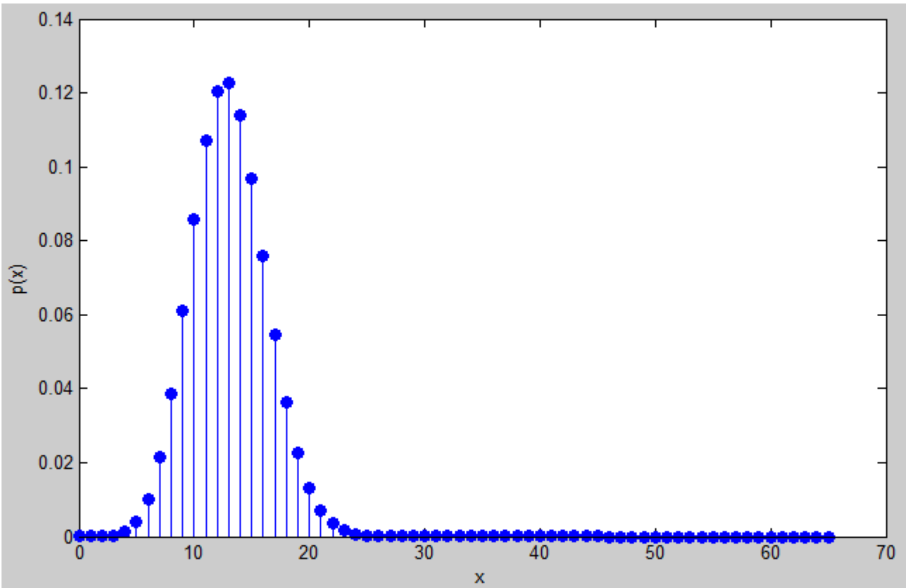
$\mu_3 = -1{,}3387$

$\mu_4 = 11{,}4224$

$\alpha_3 = -0{,}5062$

$\alpha_4 = 0{,}1229$

Obr. 4 Pravděpodobnostní funkce náhodné veličiny $X \sim Bi(15; 0{,}85)$



$EX = 13$

$DX = 10{,}4$

$\sigma(X) = 3{,}2249$

$\mu_3 = 6{,}24$

$\mu_4 = 324{,}896$

$\alpha_3 = 0{,}1861$

$\alpha_4 = 0{,}0038$

Obr. 5 Pravděpodobnostní funkce náhodné veličiny $X \sim Bi(65; 0{,}2)$

Z obr. 4 a 5 a z číselných charakteristik je vidět, že i pro různé parametry mají oba příklady podobné střední hodnoty. Pro $X \sim Bi(65; 0{,}2)$ je však výrazně větší rozptyl a směrodatná odchylka. Špičatost je rovna téměř nule, což je v grafu dobře vidět. U $X \sim Bi(15; 0{,}85)$ stojí za zmínku výrazná záporná šikmost, zobrazená pozvolným nárůstem hodnot a strmým poklesem.

2.3 Poissonovo rozdělení

Jelikož také při výpočtu hodnot pravděpodobnostní funkce binomického rozdělení je potřeba počítat kombinační číslo, a tedy faktoriály, nahrazuje se binomické rozdělení Poissonovým. To je však možné pouze za podmínek uvedených v kapitole 2.4.

Poissonovo rozdělení je diskrétní rozdělení pravděpodobnosti. Nechť náhodná veličina X má obor hodnot $M = \{0,1,2,\dots\}$ a pravděpodobnostní funkci určenou vztahem:

$$p_{Po}(x) = P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x \in M$$

$$p_{Po}(x) = P(X = x) = 0, \qquad x \notin M$$

Pak říkáme, že náhodná veličina X má Poissonovo rozdělení pravděpodobnosti s parametrem $\lambda > 0$. Označení: $X \sim Po(\lambda)$.

Věta: Nechť náhodná veličina X má Poissonovo rozdělení pravděpodobnosti, pak číselné charakteristiky jsou dány vztahy:

- střední hodnota $EX = \lambda$
- rozptyl $DX = \lambda$
- směrodatná odchylka $\sigma(X) = \sqrt{\lambda}$
- 3. centrální moment $\mu_3 = \lambda$
- 4. centrální moment $\mu_4 = \lambda + 3\lambda^2$
- šikmost $\alpha_3 = \frac{1}{\sqrt{\lambda}}$
- špičatost $\alpha_4 = \frac{1}{\lambda}$

Důkaz: Přeznačme $p_{Po}(x)$ na $p(x)$. I zde je ukázáno obou možných způsobů výpočtu obecných momentů EX^n. Hodnoty EX a EX^2 jsou určeny podle vztahu

$$EX^n = \sum_{x \in M} x^n p(x),$$

hodnoty EX^3 a EX^4 pak pomocí momentové vytvořující funkce.

$$EX = \sum_{x \in M} xp(x) = \sum_{x=0}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!} = \sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^x}{(x-1)!} = \lambda\sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^{x-1}}{(x-1)!} = \lambda$$

$$EX^2 = \sum_{x \in M} x^2 p(x) = \sum_{x=0}^{\infty}(x(x-1)+x)p(x) = \sum_{x=1}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} + \sum_{x=1}^{\infty} xp(x) =$$

$$= \sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^x}{(x-2)!} + EX = \lambda^2\sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^{x-2}}{(x-2)!} + \lambda = \lambda^2 + \lambda$$

$$DX = EX^2 - E^2X = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$\sigma(X) = \sqrt{DX} = \sqrt{\lambda}$$

Pro výpočet 3. a 4. centrálního momentu je potřeba nejdříve určit EX^3 a EX^4. Zde se využívá momentové vytvořující funkce, která má pro Poissonovo rozdělení podle [13] tvar

$$M_X(t) = e^{\lambda(e^t-1)}.$$

$$EX^3 = \frac{d^3 M_X}{dt^3}(0) = \lambda + 3\lambda^2 + \lambda^3$$

$$EX^4 = \frac{d^4 M_X}{dt^4}(0) = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4$$

$$\mu_3 = E(X - EX)^3 = EX^3 - 3EX^2EX + 2E^3X = (\lambda + 3\lambda^2 + \lambda^3) - 3(\lambda^2 + \lambda)\lambda + 2\lambda^3 =$$
$$= \lambda + 3\lambda^2 + \lambda^3 - 3\lambda^3 - 3\lambda^2 + 2\lambda^3 = \lambda$$

$$\mu_4 = E(X - EX)^4 = EX^4 - 4EX^3EX + 6EX^2E^2X - 3E^4X =$$
$$= (\lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4) - 4(\lambda + 3\lambda^2 + \lambda^3)\lambda + 6(\lambda^2 + \lambda)\lambda^2 - 3\lambda^4 =$$
$$= \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4 - 4\lambda^2 - 12\lambda^3 - 4\lambda^4 + 6\lambda^4 + 6\lambda^3 - 3\lambda^4 = \lambda + 3\lambda^2$$

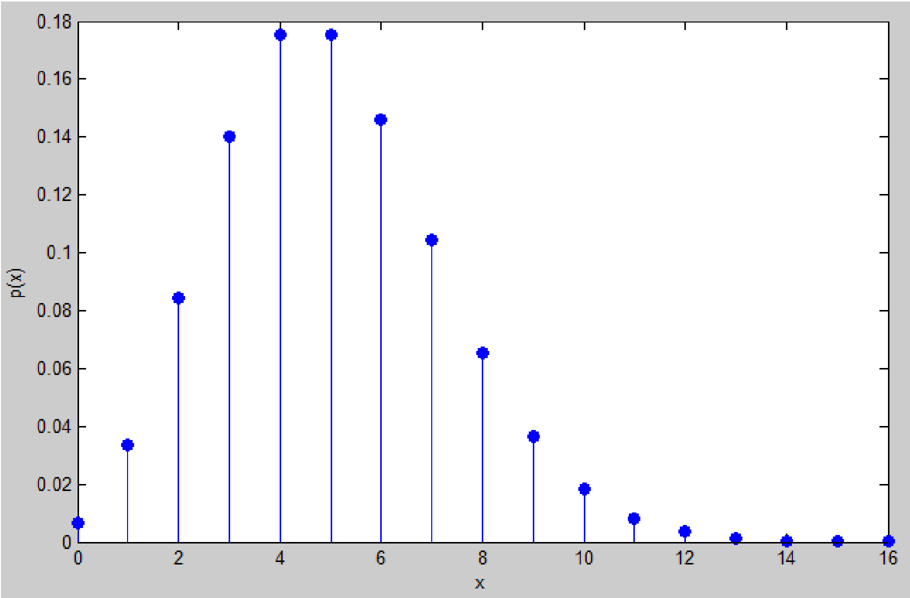
$$\alpha_3 = \frac{\mu_3}{\sigma(X)^3} = \frac{\lambda}{\sqrt{\lambda}^3} = \frac{1}{\sqrt{\lambda}}$$

$$\alpha_4 = \frac{\mu_4}{\sigma(X)^4} - 3 = \frac{\lambda + 3\lambda^2}{\sqrt{\lambda}^4} - 3 = \frac{\lambda(1 + 3\lambda)}{\lambda^2} - 3 = \frac{1}{\lambda} + 3 - 3 = \frac{1}{\lambda}$$

Tvar pravděpodobnostní funkce pro Poissonovo rozdělení s různým parametrem je dále graficky znázorněn na obr. 6 a 7. Pro oba uváděné příklady jsou vypočítány číselné charakteristiky, které jsou následně porovnány.

Z obou grafů je dobře vidět že střední hodnota má vždy hodnotu λ. Se vzrůstající hodnotou parametru roste také rozptyl a směrodatná odchylka. Pro $X{\sim}Po(5)$ se projevuje kladná šikmost tím, že hodnoty rychle narůstají a pozvolněji klesají. Pro $X{\sim}Po(35)$ vychází téměř nulová špičatost, což se projevuje tím, že tvar pravděpodobnostní funkce není zploštělý ani špičatý. Pro rostoucí parametr λ šikmost a špičatost u Poissonova rozdělení klesá k nule.

Grafické znázornění pravděpodobnostní funkce s vypočtenými číselnými charakteristikami:



$EX = 5$

$DX = 5$

$\sigma(X) = 2{,}2361$

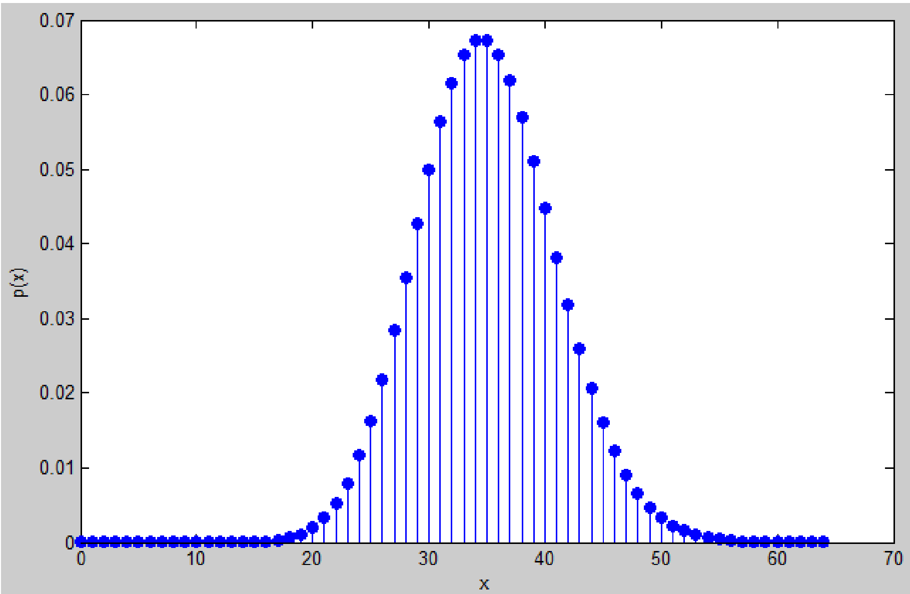
$\mu_3 = 5$

$\mu_4 = 80$

$\alpha_3 = 0{,}4472$

$\alpha_4 = 0{,}2$

Obr. 6 Pravděpodobnostní funkce náhodné veličiny $X \sim Po(5)$



$EX = 35$

$DX = 35$

$\sigma(X) = 5{,}9161$

$\mu_3 = 35$

$\mu_4 = 3710$

$\alpha_3 = 0{,}1690$

$\alpha_4 = 0{,}0286$

Obr. 7 Pravděpodobnostní funkce náhodné veličiny $X \sim Po(35)$

2.4 Vztahy mezi hypergeometrickým, binomickým a Poissonovým rozdělením

Při zkoumání asymptotického chování diskrétních rozdělení pravděpodobnosti studovaných v předchozích odstavcích, lze zjistit, že je možné některá rozdělení aproximovat jinými. V tomto odstavci bude ukázáno, jaká rozdělení lze aproximovat a za jakých podmínek je toto nahrazení považováno za dostatečně přesné.

Věta: Nechť $X \sim Hg(N, K, n)$ s pravděpodobnostní funkcí $p_{Hg}(x)$, kde $x \in \mathrm{M} = \{\max\{0, \mathrm{n} + \mathrm{K} - \mathrm{N}\}, \max\{0, \mathrm{n} + \mathrm{K} - \mathrm{N}\} + 1, \dots, \min\{\mathrm{n}, \mathrm{K}\}\}$. Když pro $N \to \infty$ platí, že

$$\lim_{N \to \infty} \frac{K}{N} = \theta \,,$$

pak platí

$$\lim_{N \to \infty} p_{Hg}(x) = p_{Bi}(x) \,.$$

Pozn.: Podle autorů (viz. [6], [7]) lze v praxi tuto aproximaci provést při splnění podmínky:

$$\frac{n}{N} < 0{,}1 \,. \tag{1}$$

Důkaz: Aproximace hypergeometrického rozdělení binomickým:

$$\lim_{N \to \infty} p_{Hg}(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} = \lim_{N \to \infty} \frac{\dfrac{K!}{x!\,(K-x)!} \cdot \dfrac{(N-K)!}{(n-x)!\,(N-K-(n-x))!}}{\dfrac{N!}{n!\,(N-n)!}} =$$

$$= \lim_{N \to \infty} \frac{n!}{x!\,(n-x)!} \cdot \frac{K(K-1)\dots(K-x+1) \cdot (N-K)(N-K-1) \cdot \dots \cdot (N-K-(n-x)+1)}{N(N-1) \cdot \dots \cdot (N-n+1)} =$$

$$= \lim_{N \to \infty} \binom{n}{x} \frac{\prod_{i=0}^{x-1}(K-i) \prod_{i=0}^{n-x-1}(N-K-i)}{\prod_{i=0}^{n-1}(N-i)} \frac{\frac{1}{N^n}}{\frac{1}{N^n}} =$$

$$= \lim_{N \to \infty} \binom{n}{x} \frac{\prod_{i=0}^{x-1}\left(\frac{K}{N} - \frac{i}{N}\right) \prod_{i=0}^{n-x-1}(1 - \frac{K}{N} - \frac{i}{N})}{\prod_{i=0}^{n-1}(1 - \frac{i}{N})} = \binom{n}{x} \theta^x (1-\theta)^{n-x} = p_{Bi}(x)$$

Věta: Nechť $X \sim Bi(n, \theta)$ s pravděpodobnostní funkcí $p_{Bi}(x)$, kde $x \in \mathrm{M} = \{0,1,2, \dots, \mathrm{n}\}$. Když pro $n \to \infty$ platí, že

$$\lim_{n \to \infty} n\theta = \lambda \,,$$

pak platí

$$\lim_{n \to \infty} p_{Bi}(x) = p_{Po}(x) \,.$$

Pozn.: Podle autorů (viz. [6], [7]) lze v praxi tuto aproximaci provést při splnění podmínek:

$$n > 30, \;\; \theta < 0{,}1 \,. \tag{2}$$

Důkaz: Aproximace binomického rozdělení Poissonovým:

Nejprve je vypočítána limita, která bude následně využita.

$$\lim_{n \to \infty} \left(\frac{n!}{n^x (n-x)!}\right) = \lim_{n \to \infty} \frac{n(n-1) \cdot \dots \cdot (n-x+1)}{n^x} = \lim_{n \to \infty} \frac{n(n-1) \cdot \dots \cdot (n-x+1)}{n \cdot n \cdot \dots \cdot n} =$$

$$= \lim_{n\to\infty} 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \ldots \cdot (1 - \frac{x-1}{n}) = 1 \cdot (1-0) \cdot \ldots \cdot (1-0) = 1$$

Dále je tedy možno psát

$$\lim_{n\to\infty} p_{Bi}(x) = \lim_{n\to\infty} \binom{n}{x} \theta^x (1-\theta)^{n-x} = \lim_{n\to\infty} \frac{n!}{x!\,(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} =$$

$$= \lim_{n\to\infty} \left(\frac{n!}{n^x (n-x)!}\right) \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} = 1 \cdot \frac{\lambda^x}{x!} e^{-\lambda} \cdot 1 = e^{-\lambda} \frac{\lambda^x}{x!} = p_{Po}(x)$$

Lze-li aproximovat hypergeometrické rozdělení binomickým a binomické rozdělení Poissonovým, pak lze také aproximovat hypergeometrické rozdělení Poissonovým. Tato aproximace se využívá zejména pro snížení výpočetní náročnosti a podle výše zmíněných autorů ji lze provést při splnění podmínek

$$\frac{n}{N} < 0{,}1, \quad \frac{K}{N} < 0{,}1, \quad n > 30\,. \tag{3}$$

Pozn.: Zde odvozených vztahů se využívá při konstrukci přejímacích plánů u statistických přejímek srovnáváním.

K ověření předepisovaných podmínek a ke grafickému zobrazení jednotlivých aproximací je v Matlabu vytvořen program (viz. Přílohy - Program1, popř. kapitola 2.6), ve kterém lze tyto aproximace simulovat. Dále jsou na obr. 8 a 9 uvedeny grafické výstupy z tohoto programu. Jsou na nich vidět pravděpodobnostní funkce hypergeometrického, binomického a Poissonova rozdělení pro případ, kdy nesplňují (obr. 8) popř. splňují (obr. 9) podmínky

$$\frac{n}{N} < 0{,}1, \quad \frac{K}{N} < 0{,}1, \quad n > 30.$$

Jak moc se k sobě jednotlivá rozdělení přibližují, je počítáno pomocí euklidovské normy vektorů rozdílů hodnot pravděpodobnostních funkcí. Nechť $\boldsymbol{y_{Hg}}$ je vektor hodnot pravděpodobnostní funkce hypergeometrického rozdělení, $\boldsymbol{y_{Bi}}$ je vektor hodnot pravděpodobnostní funkce binomického rozdělení a $\boldsymbol{y_{Po}}$ je vektor hodnot pravděpodobnostní funkce Poissonova rozdělení. Odchylky mezi jednotlivými pravděpodobnostními funkcemi (Hg značí hypergeometrické rozdělení, Bi – binomické rozdělení a Po - Poissonovo rozdělení) jsou počítány následujícím způsobem:

$$\rho(Hg, Bi) = \left\|\boldsymbol{y_{Hg}} - \boldsymbol{y_{Bi}}\right\|_2$$

$$\rho(Bi, Po) = \|\boldsymbol{y_{Bi}} - \boldsymbol{y_{Po}}\|_2$$

$$\rho(Hg, Po) = \left\|\boldsymbol{y_{Hg}} - \boldsymbol{y_{Po}}\right\|_2$$

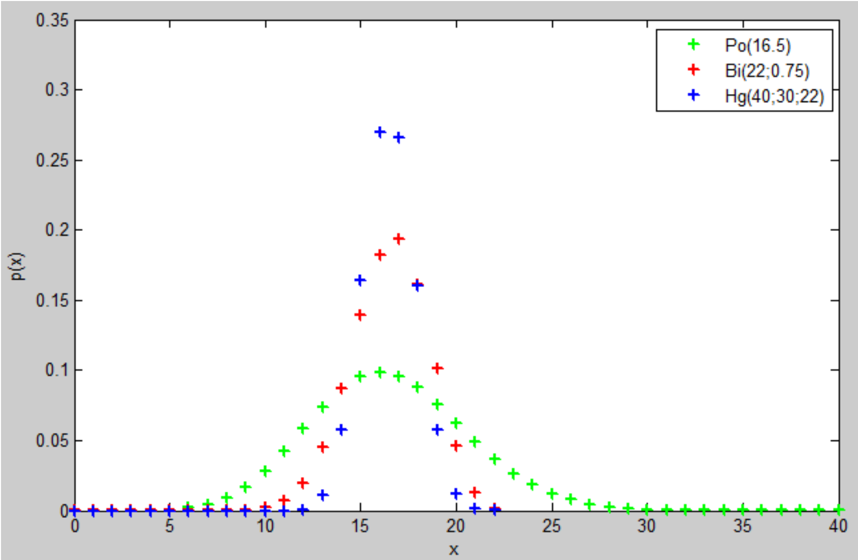
Pozn.: Ve vytvořeném programu Program1 (viz. přílohy) je voleno označení:

$$\rho(Hg, Bi) = normaHgBi$$

$$\rho(Bi, Po) = normaBiPo$$

$$\rho(Hg, Po) = normaHgPo\,.$$

Grafické znázornění konvergencí:



$$\rho(Hg, Bi) = 0{,}1384$$
$$\rho(Bi, Po) = 0{,}1789$$
$$\rho(Hg, Po) = 0{,}2932$$
$$\frac{n}{N} = 0{,}55 > 0{,}1$$
$$\frac{K}{N} = 0{,}75 > 0{,}1$$
$$n = 22 < 30$$

Obr. 8 Průběhy pravděpodobnostních funkcí při nesplněných podmínkách pro aproximaci rozděleni

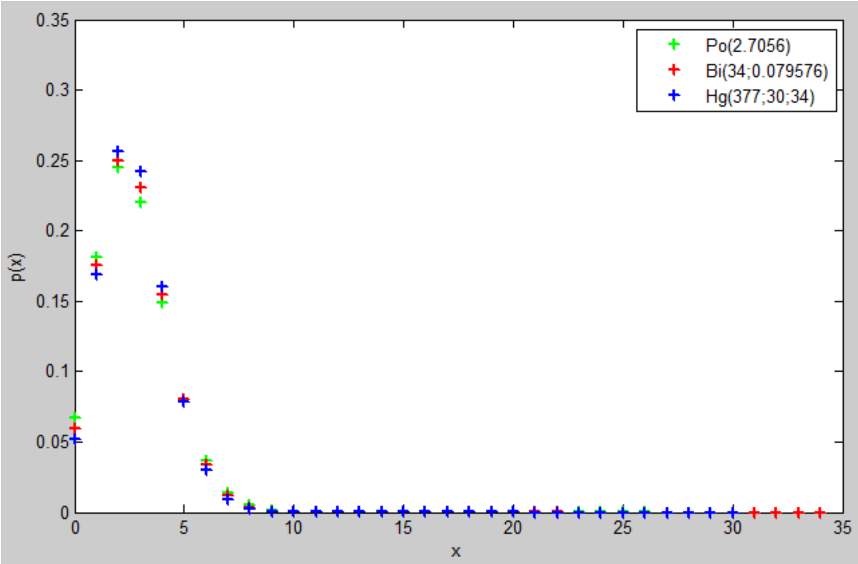
K obr. 8 je doplněna tabulka pro porovnání hodnot pravděpodobnostních funkcí pro $x \in \langle 11, 20 \rangle$ při nesplněných podmínkách pro aproximaci rozdělení.

Tab.1: Hodnoty pravděpodobnostních funkcí při nesplněných podmínkách pro aproximaci rozdělení

x / p(x)	11	12	13	14	15	16	17	18	19	20
y_{Hg}	0,0000	0,0008	0,0106	0,0577	0,1642	0,2693	0,2662	0,1602	0,0578	0,0119
y_{Bi}	0,0071	0,0195	0,0451	0,0869	0,1391	0,1826	0,1933	0,1611	0,1017	0,0458
y_{Po}	0,0422	0,0580	0,0736	0,0868	0,0955	0,0985	0,0956	0,0876	0,0761	0,0628

Na dalším obrázku (obr. 9) je ukázáno, jak se změní tvary pravděpodobnostních funkcí, když jsou podmínky pro aproximaci jednotlivých rozdělení splněny. K obr. 9 je opět doplněna tabulka pro porovnání hodnot pravděpodobnostních funkcí pro $x \in \langle 0{,}9 \rangle$, tentokrát už při splněných podmínkách pro aproximaci rozdělení.

Rozdíl v obou obrázcích je značný. Na obr. 8 mají pravděpodobnostní funkce zcela odlišný tvar, což potvrzuje i to, že ani jedna ze tří podmínek pro možnou aproximaci není splněna. Na obr. 9 jsou už podmínky splněny a tvary všech tří pravděpodobnostních funkcí jsou dosti podobné. Pro tento případ také výrazně klesla norma vektoru rozdílů hodnot jednotlivých pravděpodobnostních funkcí.



$\rho(Hg, Bi) = 0{,}0184$
$\rho(Bi, Po) = 0{,}0161$
$\rho(Hg, Po) = 0{,}0345$
$\frac{n}{N} = 0{,}0902 < 0{,}1$
$\frac{K}{N} = 0{,}0796 < 0{,}1$
$n = 34 > 30$

Obr. 9 Průběhy pravděpodobnostních funkcí při splněných podmínkách pro aproximaci rozděleni

Tab. 2: Hodnoty pravděpodobnostních funkcí při splněných podmínkách pro aproximaci rozdělení

x / p(x)	0	1	2	3	4	5	6	7	8	9
y_{Hg}	0,0520	0,1689	0,2566	0,2425	0,1601	0,0785	0,0297	0,0089	0,0022	0,0004
y_{Bi}	0,0596	0,1753	0,2501	0,2307	0,1545	0,0802	0,0335	0,0116	0,0034	0,0008
y_{Po}	0,0668	0,1808	0,2446	0,2206	0,1492	0,0807	0,0364	0,0141	0,0048	0,0014

Na základě provedených simulací lze usoudit, že uvedené podmínky

$$\frac{n}{N} < 0{,}1, \quad \frac{K}{N} < 0{,}1, \quad n > 30$$

lze považovat za vhodné. Představa o chybě je vidět z obr. 8 a 9, popř. z tab. 1 a 2.

2.5 Distribuční funkce

Distribuční funkce hypergeometrického, binomického a Poissonova rozdělení budou dále označovány následujícím způsobem:

$$F_{Hg}(x; N, K, n) = \sum_{i=0}^{x} \frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}}$$

$$F_{Bi}(x; n, \theta) = \sum_{i=0}^{x} \binom{n}{i} \theta^i (1-\theta)^{n-i}$$

$$F_{Po}(x; \lambda) = \sum_{i=0}^{x} \frac{e^{-\lambda}(\lambda)^i}{i!}.$$

Grafické zobrazení jednotlivých distribučních funkcí si lze prohlédnout ve vytvořeném programu v Matlabu (viz. přílohy Program1, popř. kapitola 2.6). Zde je možnost vykreslení distribuční funkce pro dané diskrétní rozdělení nebo ukázky simulace aproximace jednotlivých distribučních funkcí. Dále jsou na obr. 10 a 11 uvedeny grafické výstupy z tohoto programu. Jsou na nich vidět distribuční funkce hypergeometrického, binomického a Poissonova rozdělení pro případ, kdy nesplňují (obr. 10) popř. splňují (obr. 11) podmínky

$$\frac{n}{N} < 0{,}1, \;\; \frac{K}{N} < 0{,}1, \;\; n > 30\,.$$

Jak moc se k sobě jednotlivá rozdělení přibližují, je počítáno pomocí euklidovské normy vektorů rozdílů hodnot distribučních funkcí. Nechť $\boldsymbol{v_{Hg}}$ je vektor hodnot distribuční funkce hypergeometrického rozdělení, $\boldsymbol{v_{Bi}}$ je vektor hodnot distribuční funkce binomického rozdělení a $\boldsymbol{v_{Po}}$ je vektor hodnot distribuční funkce Poissonova rozdělení. Odchylky mezi jednotlivými distribučními funkcemi (Hg značí hypergeometrické, Bi - binomické a Po - Poissonovo rozdělení) jsou počítány následujícím způsobem:

$$\rho(Hg, Bi) = \left\|\boldsymbol{v_{Hg}} - \boldsymbol{v_{Bi}}\right\|_2$$

$$\rho(Bi, Po) = \|\boldsymbol{v_{Bi}} - \boldsymbol{v_{Po}}\|_2$$

$$\rho(Hg, Po) = \left\|\boldsymbol{v_{Hg}} - \boldsymbol{v_{Po}}\right\|_2$$

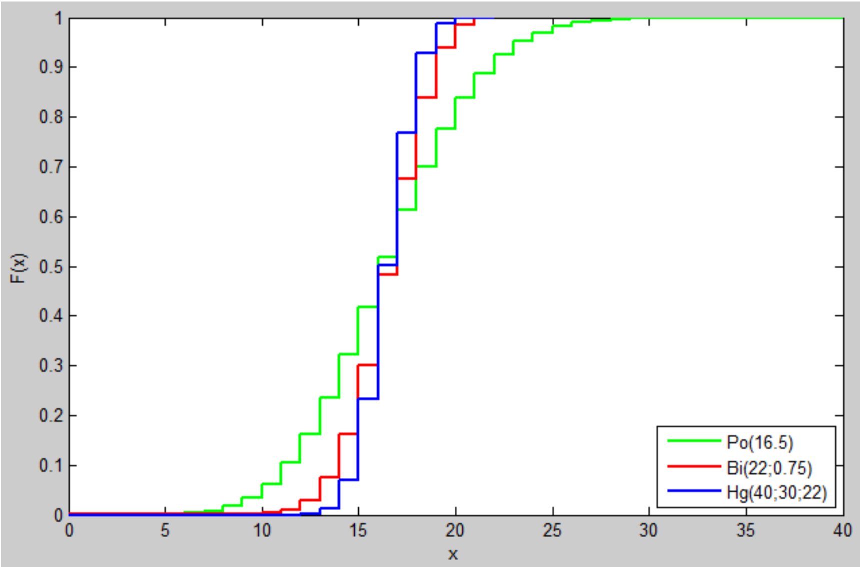
Pozn.: Ve vytvořeném programu Program1 (viz. přílohy) je voleno označení:

$$\rho(Hg, Bi) = normaHgBi$$

$$\rho(Bi, Po) = normaBiPo$$

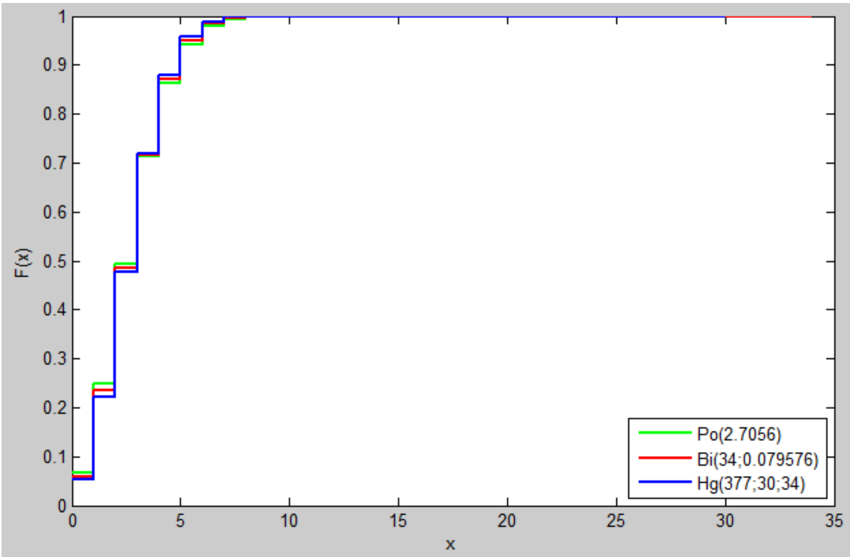
$$\rho(Hg, Po) = normaHgPo\,.$$

Grafické znázornění konvergencí:



$$\rho(Hg, Bi) = 0{,}1942$$
$$\rho(Bi, Po) = 0{,}4329$$
$$\rho(Hg, Po) = 0{,}5973$$
$$\frac{n}{N} = 0{,}55 > 0{,}1$$
$$\frac{K}{N} = 0{,}75 > 0{,}1$$
$$n = 22 < 30$$

Obr. 10 Průběhy distribučních funkcí při nesplněných podmínkách pro aproximaci rozdělení



$$\rho(Hg, Bi) = 0{,}0228$$
$$\rho(Bi, Po) = 0{,}0206$$
$$\rho(Hg, Po) = 0{,}0434$$
$$\frac{n}{N} = 0{,}0902 < 0{,}1$$
$$\frac{K}{N} = 0{,}0796 < 0{,}1$$
$$n = 34 > 30$$

Obr. 11 Průběhy distribučních funkcí při splněných podmínkách pro aproximaci rozdělení

Stejně jako u konvergencí pravděpodobnostních funkcí, tak i zde je rozdíl v průběhu distribučních funkcí značný. Na obr. 10 mají distribuční funkce zcela odlišný tvar, což potvrzuje i to, že ani jedna ze tří podmínek není splněna. Na obr. 11 jsou už podmínky splněny a tvary všech tří distribučních funkcí jsou dosti podobné. Pro tento případ také výrazně klesla norma vektorů rozdílů hodnot jednotlivých distribučních funkcí a lze je tedy navzájem aproximovat.

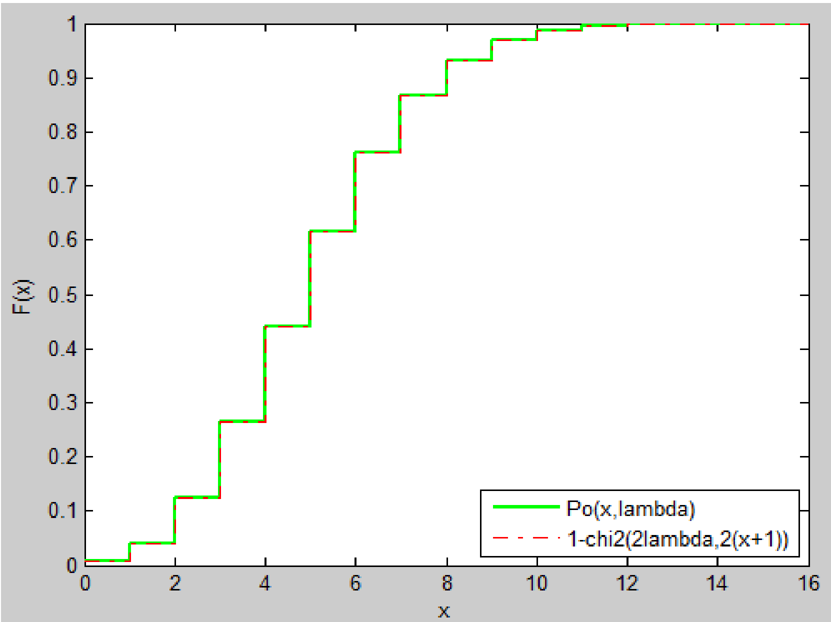
Při konstrukcích přejímacích plánů je třeba distribuční funkce počítat a dále se využívá jejich vzájemných aproximací. Kromě toho se využívá vztahu, který dává do souvislosti distribuční funkci Poissonova a Pearsonova χ^2 rozdělení. Ten objasňuje následující věta uvedená v [5].

Věta: Nechť náhodná veličina X má rozdělení Po(λ) a náhodná veličina Y má rozdělení χ^2 o 2(x+1) stupních volnosti. Pak platí

$$F_{Po}(x;\lambda) = \sum_{i=0}^{x} \frac{\lambda^i}{i!} e^{-\lambda} = \frac{1}{2^{x+1}\Gamma(x+1)} \int_{2\lambda}^{\infty} e^{-\frac{y}{2}} y^x dy = P(Y > 2\lambda) = 1 - F_{\chi^2}(2\lambda; 2(x+1)) \,. \tag{4}$$

Důkaz této věty lze najít v [5] na straně 141.

Platnost vztahu $F_{Po}(x;\lambda) = 1 - F_{\chi^2}(2\lambda; 2(x+1))$ je také ověřena numericky. Na obr. 12 jsou vykresleny distribuční funkce Poissonova rozdělení a funkce $1 - F_{\chi^2}(2\lambda; 2(x+1))$. Je vidět, že průběhy obou funkcí jsou totožné.



Obr. 12 Grafický znázornění vztahu $F_{Po}(x;\lambda) = 1 - F_{\chi^2}\big(2\lambda; 2(x+1)\big)$ pro $\lambda = 5$.

2.6 Počítačová implementace

Pro grafickou demonstraci pravděpodobnostních a distribučních funkcí diskrétních rozdělení, výpočty jejich číselných charakteristik a simulací konvergencí z kapitoly 2.4 byl vytvořen počítačový program Program1 (viz. přílohy).

Program je naprogramován v Matlabu R2010a metodou Switched board programming. V Matlabu je možné v toolboxu Statistica najít funkce pro výpočet pravděpodobnostní funkce (pdf – probability density function) a distribuční funkce (cdf – cumulative distribution function). Při programování byly použity následující funkce:

hygepdf(x,N,K,n) - výpočet hodnot pravděpodobnostní funkce hypergeometrického rozdělení v bodech určených vektorem x

binopdf(x,n,θ) - výpočet hodnot pravděpodobnostní funkce binomického rozdělení v bodech určených vektorem x

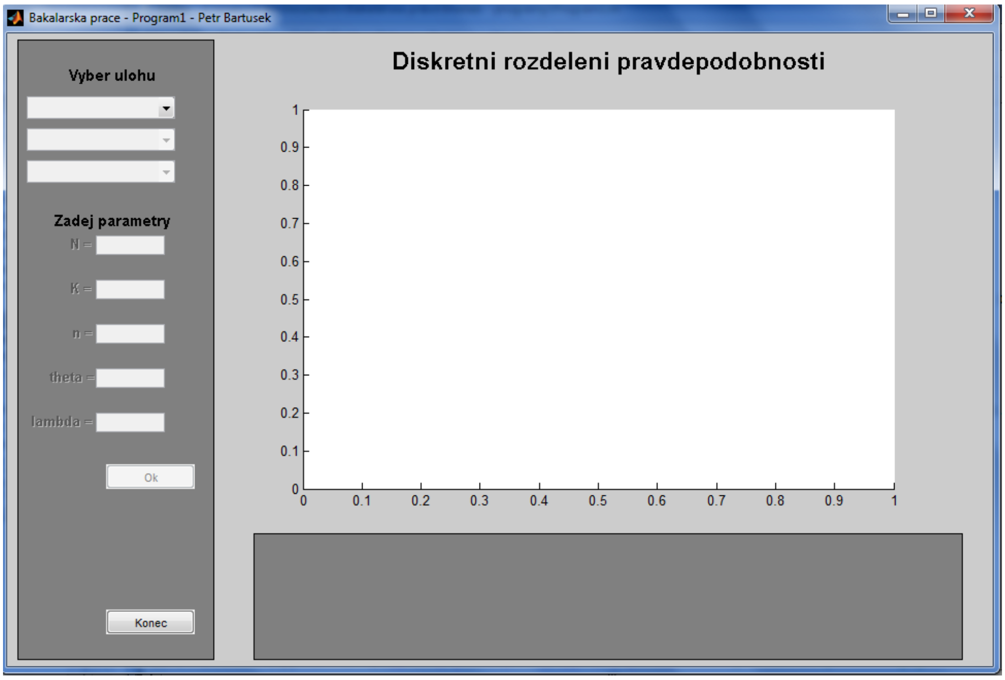
poisspdf(x,λ) - výpočet hodnot pravděpodobnostní funkce Poissonova rozdělení v bodech určených vektorem x

hygecdf(x,N,K,n) - výpočet hodnot distribuční funkce hypergeometrického rozdělení v bodech určených vektorem x

binocdf(x,n,θ) - výpočet hodnot distribuční funkce binomického rozdělení v bodech určených vektorem x

poisscdf(x,λ) - výpočet hodnot distribuční funkce Poissonova rozdělení v bodech určených vektorem x

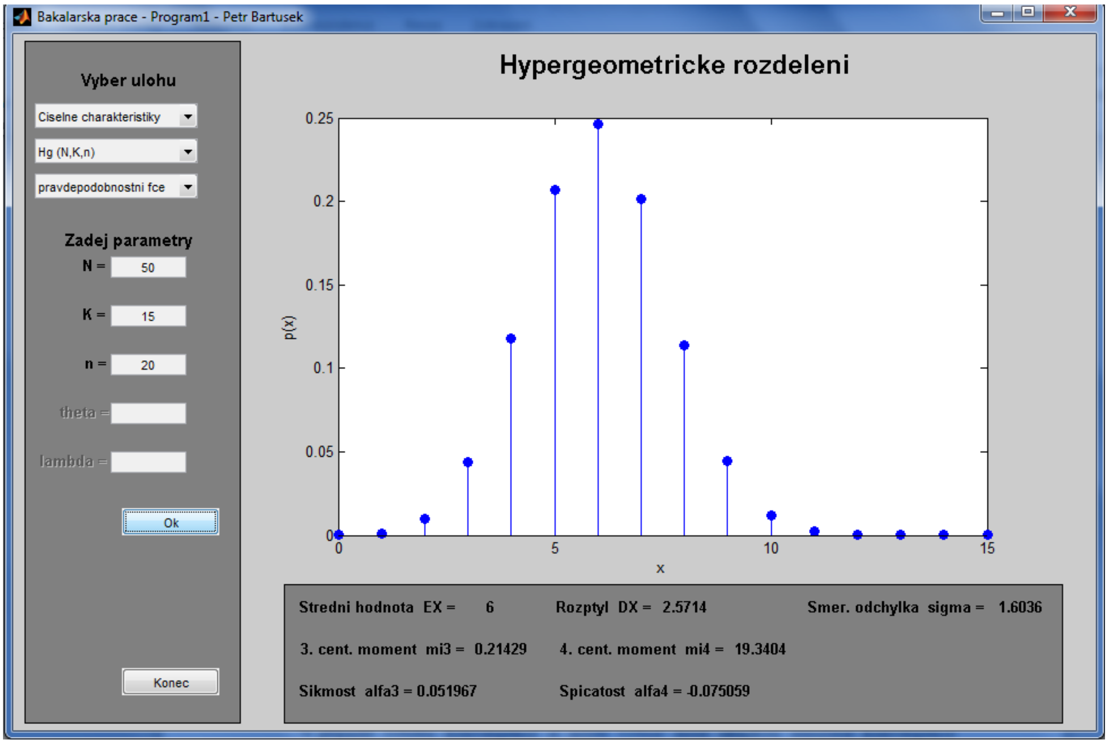
Při spuštění se uživateli objeví grafické okno. V levé části se nacházejí tři rozbalovací menu pro výběr dané úlohy a edit okna pro zadání vstupních parametrů. V pravé horní části je místo pro grafické zobrazení zadané úlohy, tj. pro vykreslování pravděpodobnostních a distribučních funkcí. V pravé dolní části jsou pak výsledky výpočtu.



Obr. 13 Program1 – grafické rozhraní

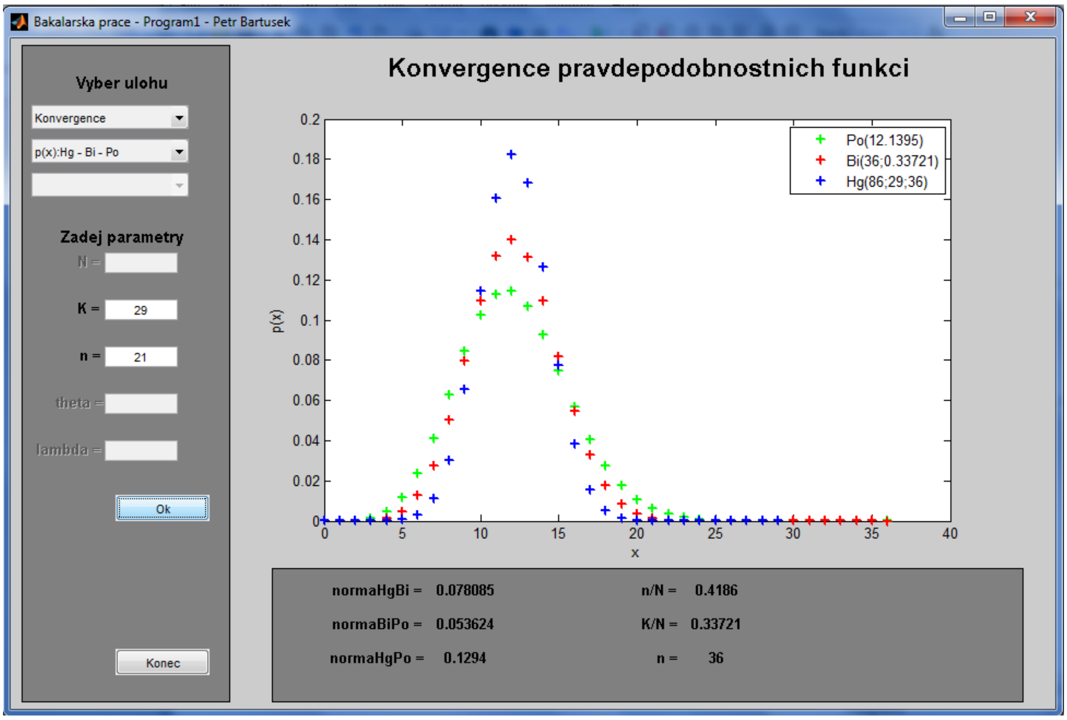
V levé části ve třech rozbalovacích menu si uživatel může vybrat, zda si chce prohlédnout ukázku s výpočtem číselných charakteristik nebo simulace konvergence.

V případě číselných charakteristik musí dále vybrat jedno ze tří nabízených diskrétních rozdělení a typ grafu (pravděpodobnostní nebo distribuční funkce), který chce zobrazit. Po vyplnění požadovaných parametrů a stisknutí tlačítka Ok, je vykreslen graf a spočteny číselné charakteristiky. Při zadávání parametrů musí uživatel vyplnit všechny viditelné kolonky, ty ostatní jsou pro uživatele v daný moment nedostupné. Do edit oken pro parametry N, K a n je nutné zadat kladné celé číslo a při výběru hypergeometrického rozdělení musí platit $N \geq K \ \wedge \ N \geq n$. Lambda i theta jsou reálná čísla ($lambda > 0 \ a \ theta \in \langle 0; 1 \rangle$). Při zadávání těchto hodnot je nutné dát pozor zejména na to, že se místo desetinné čárky píše tečka.



Obr. 14 Program1 – ukázka vykreslení pravděpodobnostní funkce a znázornění vypočtených číselných charakteristik hypergeometrického rozdělení pro zadané parametry N, K, n

V případě výběru konvergence, si uživatel může vybrat mezi ukázkou simulace konvergencí pravděpodobnostních nebo distribučních funkcí hypergeometrického, binomického a Poissonova rozdělení. Po vyplnění požadovaných parametrů a stisknutí tlačítka Ok, jsou simulace spuštěny. Souběžně jsou počítány některé ukazatele, např. norma vektoru rozdílů hodnot pravděpodobnostních (popř. distribučních) funkcí a ověřovány podmínky pro možnou aproximaci rozdělení z kapitoly 2.4.



Obr. 15 Program1 – ukázka vykreslení a znázornění vypočtených ukazatelů pro konvergenci pravděpodobnostních funkcí

3. Statistické přejímky

3.1 Úvod do statistických přejímek

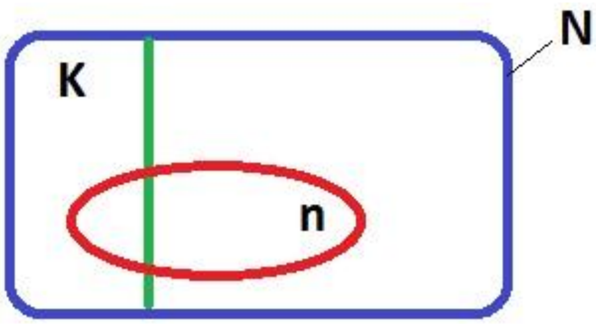
Při výrobě průmyslových součástek je potřeba kontrolovat jejich jakost. Pokud odběratel odebírá dodávky nových výrobků od dodavatele, provádí jejich kontrolu (přejímací kontrola), aby se ujistil, že procento zmetků nepřesahuje přípustnou mez. *Statistické přejímky* jsou postupy, které řeší, zda bude takováto dodávka přijata, či zamítnuta. Přejímací postupy by měly také dávat dodavateli i odběrateli určité záruky a zajistit, aby provedení kontroly bylo co nejekonomičtější, tzn. náklady potřebné na kontrolu výrobků byly co nejmenší. Zárukou se myslí to, aby odběratel nepřijímal dodávky s nevyhovující jakostí a naopak, aby se dodavateli nevracely dodávky s přijatelnou jakostí.

Nechť N je celkový počet výrobků v dodávce, z nichž K má určitou vlastnost (např. jsou vadné, mají nevyhovující jakost). Přejímací kontroly se podle rozsahu kontrolovaných výrobků dělí na:

- stoprocentní
- výběrové

Stoprocentní kontrola probíhá tak, že se překontroluje všech N výrobků. Pokud se mezi nimi zjistí K vadných kusů, pak číslo $p = K/N$ udává podíl vadných výrobků v celé dodávce. Stoprocentní kontroly mají však několik nevýhod, např. vysoké náklady, časovou náročnost nebo někdy i poničení výrobku (destruktivní zkoušky, změna jakosti).

Kvůli těmto nevýhodám stoprocentní kontroly se u dodávek provádí výběrová přejímací kontrola jakosti. U této kontroly se nekontroluje všech N výrobku, ale jen určitý počet n z celé dodávky. Náhodný výběr u výběrové přejímací kontroly jakosti odpovídá urnovému schématu. Proto se také při výpočtech vychází z hypergeometrického rozdělení pravděpodobnosti, které je možno, po splnění podmínek (1) a (2) uvedených v kapitole 2.4, aproximovat rozdělením binomickým, popř. Poissonovým.



Obr. 16: Grafické znázornění náhodného výběru n výrobků z dodávky o N výrobcích, z nichž K je vadných

3.2 Rozdělení statistických přejímek

Statistické přejímky lze dělit podle nejrůznějších kritérií:

1. Kontrola jakosti výrobků
   - statistické přejímky srovnáváním
   - statistické přejímky měřením

   Při statistické přejímce srovnáváním nás nezajímá konkrétní jakostní znak, ale pouze to, zda je nebo není daný výrobek vadný. To se provádí pomocí měrek, kalibrů, úchylkometrů, podle vzhledové kontroly, srovnáváním s předepsaným vzorkem, atd.

   U statistických přejímek měřením nás zajímá konkrétní hodnota sledovaného znaku, a proto také závisejí na pravděpodobnostním rozdělení tohoto znaku.

2. Z hlediska typu poskytované záruky
   - statistické přejímky pro ověření jakosti jednotlivého souboru
   - statistické přejímky pro ověření jakosti souborů tvořících sérii

   U statistické přejímky jednotlivého souboru nelze použít žádné informace z ostatních souborů. Je nutno pracovat jen s údaji zjištěnými z dané dodávky. Kdežto u souborů tvořících sérii lze, z již zkontrolovaných souborů, získat nejrůznější informace, které se mohou využít ke stanovení nejvýhodnějšího přejímacího postupu.

3. Výběr výrobků ke kontrole
   - statistické přejímky jedním výběrem
   - statistické přejímky dvojím výběrem
   - statistické přejímky několikerým výběrem
   - statistické přejímky sekvenční

   Při statistické přejímce jedním výběrem je z celého souboru N výrobků vybíráno náhodně n výrobku ke kontrole. Poté dochází k rozhodnutí o přijetí nebo zamítnutí dodávky.

   Při statistické přejímce dvojím výběrem je nejdříve z celého souboru N výrobku vybíráno náhodně n_1 výrobků, pokud nedojde k rozhodnutí, je vybráno náhodně ze zbytku souboru dalších n_2 výrobků. Poté už definitivně dochází k rozhodnutí o přijetí nebo zamítnutí souboru.

   Při statistické přejímce několikerým výběrem se postupuje analogicky jako u dvojího výběru. Postupně se vybírá náhodně n_1, n_2, n_3, … výrobků. Pokud po žádném výběru nedojde k rozhodnutí o přijetí, dostane se až ke k-tému výběru n_k, po kterém se už definitivně rozhodne o přijetí nebo zamítnutí souboru.

   Při sekvenční statistické přejímce je vybírán z celého souboru postupně vždy jeden kus, po kterém se rozhoduje o přijetí, zamítnutí nebo pokračování v dalším výběru.

4. Postup při zamítnutí dodávky
   - statistické přejímky bezopravné
   - statistické přejímky opravné

   Pokud byla kontrolou dodávka zamítnuta, dochází u statistických přejímek opravných vždy ke stoprocentnímu překontrolování celého souboru. Dodávka se převezme až poté, co jsou vadné kusy opraveny nebo nahrazeny novými.

   U statistických přejímek bezopravných ke stoprocentní kontrole nedochází a je zkontrolován jen náhodný výběr n výrobků.

3.3 Statistické přejímky srovnáváním s jedním výběrem

U statistických přejímek srovnáváním s jedním výběrem se nekontroluje konkrétní jakostní znak, ale pouze to, zda je nebo není daný výrobek vadný. Při provádění přejímací kontroly je z celého souboru o počtu N výrobků vybíráno náhodně n výrobků ke kontrole.

Podíl vadných výrobků v dodávce je číslo $p \in \langle 0; 1 \rangle$ dané vztahem

$$p = \frac{K}{N},$$

kde K je počet vadných výrobků a N je celkový počet výrobků v dodávce. Aby se statistická přejímka mohla uskutečnit, musí se ještě před zahájením přejímací kontroly odběratel s dodavatelem společně domluvit na dvou hodnotách podílů vadných výrobků p_1 a p_2 v dodávce a hodnotách rizika dodavatele α a rizika odběratele β (viz. kapitola 3.3.1).

3.3.1 Základní pojmy

Přípustný podíl vadných výrobků v dodávce je číslo $p_1 \in \langle 0; 1 \rangle$, který chrání dodavatele před zamítáním dodávek s malým podílem vadných kusů p. Někdy se místo čísla p_1 udává hodnota P_1, která značí přípustné procento vadných výrobků v dodávce.

$$P_1 = 100p_1$$

Nepřípustný podíl vadných výrobků v dodávce je číslo $p_2 \in \langle 0; 1 \rangle$, který chrání odběratele před přijímáním dodávek s velkým podílem vadných kusů p. Někdy se místo čísla p_2 udává hodnota P_2, která značí nepřípustné procento vadných výrobků v dodávce.

$$P_2 = 100p_2$$

Přejímací plán (n, c) je dvojice čísel n a c, kde n je rozsah výběru a c je přejímací kritérium, podle nichž dojde po výběrové kontrole k rozhodnutí, zda přijmout nebo zamítnout dodávku výrobků (způsob rozhodování je vysvětlen v kapitole 3.3.2).

Přejímací (akceptační) číslo c je maximální počet vadných výrobku při přejímací kontrole n náhodně vybraných kusů z celého souboru, při němž je ještě dodávka přijata.

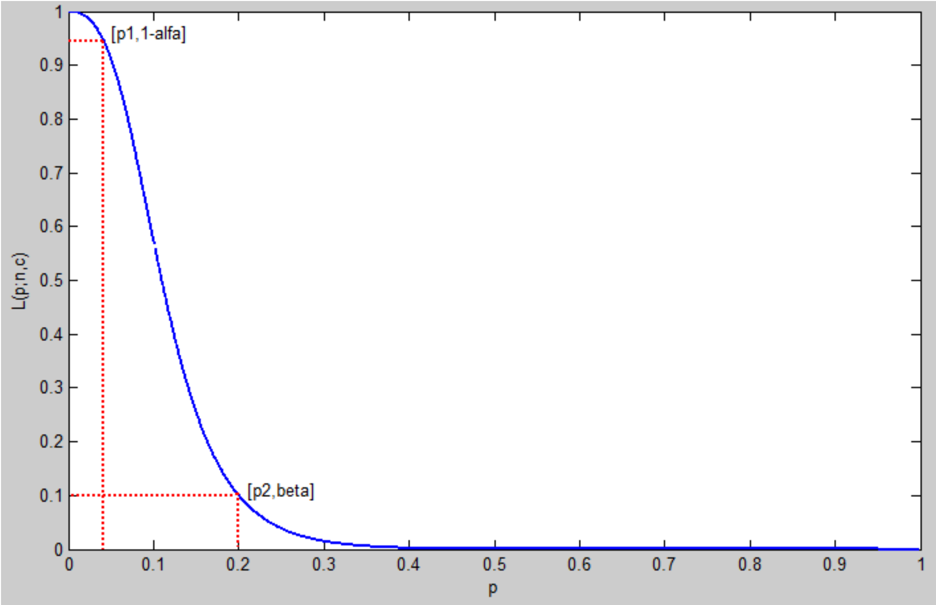
Operativní charakteristika L(p; n, c) přejímacího plánu (n, c) je funkce, udávající pravděpodobnost přijetí dodávky v závislosti na podílu vadných výrobků v dodávce p. Operativní charakteristika je klesající funkce procházející body [0,1] a [1,0]. To vychází z toho, že pravděpodobnost přijetí dodávky s malým podílem vadných kusů je vysoká a pravděpodobnost přijetí je nižší, pokud podíl vadných výrobků je větší. Pokud soubor neobsahuje žádné vadné výrobky, tedy podíl vadných je p = 0, měla by být dodávka přijímána s pravděpodobností 1 a naopak pokud soubor obsahuje samé zmetky, tedy podíl vadných je p = 1, pak by měla být dodávka přijímána s pravděpodobností 0. Operativní charakteristika vyjadřuje účinnost přejímacího plánu. Každému přejímacímu plánu odpovídá příslušná jediná operativní charakteristika. Průběh operativní charakteristiky je znázorněn na obr. 17.

Riziko dodavatele α je pravděpodobnost zamítnutí dodávky s podílem vadných kusů, který odpovídá přípustnému podílů vadných výrobků p_1 v dodávce, jenž byl určen po domluvě mezi dodavatelem a odběratelem. Nejčastěji se volí α = 0,05 nebo α = 0,01.

$$\alpha = 1 - L(p_1; n, c)$$

Riziko odběratele β je pravděpodobnost přijetí dodávky s podílem vadných kusů, který odpovídá nepřípustnému podílu vadných výrobků p_2 v dodávce, jenž byl určen po domluvě mezi dodavatelem a odběratelem. Nejčastěji se volí β = 0,05 nebo β = 0,10.

$$\beta = L(p_2; n, c)$$



Obr. 17: Operativní charakteristika se zobrazením hodnot p_1, p_2, α, β

3.3.2 Bezopravné přejímací plány (P_1, P_2)

Přejímací plán (n, c), jehož operativní charakteristika prochází dvěma body $[p_1, 1 - \alpha]$, $[p_2, \beta]$, splňuje tedy podmínky

$$L(p_1; n, c) = 1 - \alpha$$
$$L(p_2; n, c) = \beta\,, \qquad (5)$$

se nazývá přejímací plán (P_1, P_2).
U statistické přejímky jedním výběrem jsou před zahájením přejímací kontroly známy následující údaje:

N – počet výrobků v celém souboru
p_1 – přípustný podíl vadných výrobků v dodávce
p_2 – nepřípustný podíl vadných výrobků v dodávce
α – riziko dodavatele
β – riziko odběratele,

kde p_1, p_2, α, β jsou voleny po domluvě dodavatele a odběratele. Z těchto údajů se vytvoří ideální přejímací plán (n, c), podle kterého se provede přejímací kontrola.

n – rozsah výběru
c – přejímací číslo

Provede se jeden náhodný výběr rozsahu n. Tyto vybrané výrobky se překontrolují a zjistí se počet vadných výrobků. Pokud je počet vadných menší nebo roven přejímacímu číslu c, pak je celá dodávka přijata a nechává si ji odběratel. Pokud je počet vadných větší než c, pak je dodávka zamítnuta a je vrácena dodavateli. Ten musí opravit vadné kusy výrobků nebo je nahradit novými, poté může poslat dodávku opět ke kontrole.

Nechť X je počet vadných výrobků ve výběru. Postup statistické přejímky jedním výběrem je následující:

$X \leq c$ dodávka se přijímá
$X > c$ dodávka se zamítá

Dojde tedy k jednoznačnému rozhodnutí o přijetí, či zamítnutí dodávky výrobků.

3.3.3 Stanovení přejímacího plánu (n, c) pro bezopravné přejímací plány (P_1, P_2) jedním výběrem

V tomto odstavci bude popsán způsob výpočtu dvojice čísel (n, c), kde n je rozsah výběru a c je přejímací číslo.

Vychází se ze dvou podmínek (5) pro přejímací plány (P_1, P_2):

$$L(p_1; n, c) = 1 - \alpha$$
$$L(p_2; n, c) = \beta$$

Jestliže operativní charakteristika je funkce L(p; n, c), udávající pravděpodobnost přijetí dodávky v závislosti na podílu vadných výrobků v dodávce p, lze psát

$$L(p;\ n,\ c)\ =\ P(X\ \leq\ c)\ =\ \sum_{i=0}^{c} P(X = i) = F_{Hg}(c; N, Np, n)\,.$$

Dvě podmínky (5) pro přejímací plán (P_1, P_2) lze přepsat do soustavy dvou rovnic

$$F_{Hg}(c; N, Np_1, n) = \sum_{i=0}^{c} \frac{\binom{Np_1}{i}\binom{N-Np_1}{n-i}}{\binom{N}{n}} = 1 - \alpha$$

$$F_{Hg}(c; N, Np_2, n) = \sum_{i=0}^{c} \frac{\binom{Np_2}{i}\binom{N-Np_2}{n-i}}{\binom{N}{n}} = \beta\,. \qquad (6)$$

Nejdříve nalezneme přibližné řešení těchto rovnic. Postup odvození je podle [7] následující. Je-li splněna podmínka (1) uvedená v kapitole 2.4, lze hypergeometrické rozdělení aproximovat binomickým rozdělením

$$L(p; n, c) = \sum_{i=0}^{c} \frac{\binom{Np}{i}\binom{N-Np}{n-i}}{\binom{N}{n}} \doteq \sum_{i=0}^{c} \binom{n}{i} p^i (1-p)^{n-i} = F_{Bi}(c; n, p)\,.$$

Soustava rovnic (6) přejde do tvaru

$$F_{Bi}(c; n, p_1) = \sum_{i=0}^{c} \binom{n}{i} p_1^i (1-p_1)^{n-i} = 1 - \alpha$$

$$F_{Bi}(c; n, p_2) = \sum_{i=0}^{c} \binom{n}{i} p_2^i (1-p_2)^{n-i} = \beta\ . \qquad (7)$$

Dále, jsou-li splněny podmínky (2) uvedené v kapitole 2.4, lze binomické rozdělení aproximovat Poissonovým rozdělením

$$L(p; n, c) = \sum_{i=0}^{c} \binom{n}{i} p^i (1-p)^{n-i} \doteq \sum_{i=0}^{c} \frac{e^{-np}(np)^i}{i!} = F_{Po}(c; np).$$

Soustava rovnic (7) přejde do tvaru

$$F_{Po}(c; np_1) = \sum_{i=0}^{c} \frac{e^{-np_1}(np_1)^i}{i!} = 1 - \alpha$$

$$F_{Po}(c; np_1) = \sum_{i=0}^{c} \frac{e^{-np_2}(np_2)^i}{i!} = \beta \,. \qquad (8)$$

Dle vztahu (4) uvedeného v kapitole 2.5, který dává do souvislostí Poissonovo rozdělení s parametrem λ a distribuční funkci Pearsonova χ^2 rozdělení o 2(c + 1) stupních volnosti, lze soustavu (8) dále upravovat

$$F_{\chi^2}(2np_1; 2(c+1)) = \alpha$$

$$F_{\chi^2}(2np_2; 2(c+1)) = 1 - \beta \,. \qquad (9)$$

Pro kvantily Pearsonova χ^2 rozdělení o 2(c + 1) stupních volnosti platí

$$2np_1 = \chi^2_{\alpha}[2(c+1)]$$

$$2np_2 = \chi^2_{1-\beta}[2(c+1)]. \qquad (10)$$

Nakonec jednoduchou úpravou, např. porovnáním 2n, přejde soustava (10) dvou rovnic na jednu rovnici o jedné neznámé c.

$$\frac{p_2}{p_1} = \frac{\chi^2_{1-\beta}[2(c+1)]}{\chi^2_{\alpha}[2(c+1)]} \qquad (11)$$

V konečném vztahu (11) jsou hodnoty p_1 – přípustný podíl vadných výrobků v dodávce, p_2 – nepřípustný podíl vadných výrobků v dodávce, α – riziko dodavatele a β – riziko odběratele známé. Rovnice je vyřešena numericky metodou bisekce. Následně je možno dosadit c do jedné ze dvou rovnic ze soustavy (10) a vypočítat n.

Tento postup výpočtu však nelze využít, pokud je potřeba přesnější řešení. To se dostane řešením soustavy (7), ve které se vyskytují distribuční funkce binomického rozdělení. Přesné řešení se pak dostane řešením soustavy (6), ve které se vyskytují distribuční funkce hypergeometrického rozdělení.

Ve vytvořeném programu (viz přílohy Program2, popř. kapitola 3.3.4) je po zadání uživatelem možno vypočítat přejímací plán (n, c) řešením jakékoliv ze soustav rovnic (6), (7) nebo (8). Tento program je schopný vypočítat přesný výsledek pro libovolné n a c. Z důvodů výpočetní náročnosti jsou ovšem voleny pro n a c určité intervaly, což výrazně zrychluje rychlost výpočtu – voleno $n \in \langle 0; 1000 \rangle$ a $c \in \langle 0; 70 \rangle$.

Pozn.: Podle normy [3] by n mělo ležet v intervalu $\langle 0; 5955 \rangle$ a c v intervalu $\langle 0; 87 \rangle$. Horní hranice těchto intervalů a čísla blížící se těmto horním hranicím jsou však v normě zastoupeny jen minimálně. Zvolené intervaly $n \in \langle 0; 1000 \rangle$ a $c \in \langle 0; 70 \rangle$ pokrývají velkou většinu přejímacích plánů (n, c), které se v normě vyskytují.

Při řešení soustav rovnic (6), (7) nebo (8) se hledá optimální přejímací plán (n; c) takový, při kterém je euklidovská norma vektoru odchylek jednotlivých rovnic od čísel α a β minimální. Pouze u aproximace Poissonovým rozdělením se nepoužívá interval $c \in \langle 0; 70 \rangle$, ale pro přibližný odhad c_p se využije výpočet metodou bisekce a poté je pro c volen užší interval $c \in \langle c_p - 5; c_p + 5 \rangle$.

Jelikož se při odvozování vychází z hypergeometrického rozdělení, mělo by řešení soustavy (6) s distribučními funkcemi hypergeometrického rozdělení dávat přesné výsledky. Přesný výsledek řešení jsou dvě reálná čísla - n pro rozsah výběru a c pro přejímací číslo. V přejímacím plánu (n, c) ovšem mohou být obě čísla pouze nezáporná celá čísla. A právě tyto čísla se snaží najít vytvořený program (viz. přílohy Program2). Dává tedy jako výsledek nezáporná celá čísla n a c. Proto lze určit euklidovskou normu vektoru odchylek jednotlivých rovnic soustavy (6), od předem známých čísel rizika dodavatele α a rizika odběratele β, při dosazení vypočtených čísel n a c (získaných řešením rovnic (6), (7) nebo (8)). Nechť o_1 je odchylka od α a o_2 je odchylka od β:

$$o_1 = 1 - \alpha - \sum_{i=0}^{c} \frac{\binom{Np_1}{i}\binom{N-Np_1}{n-i}}{\binom{N}{n}}$$

$$o_2 = \beta - \sum_{i=0}^{c} \frac{\binom{Np_2}{i}\binom{N-Np_2}{n-i}}{\binom{N}{n}} . \qquad (12)$$

Význam dále používaných norem je následující:
1) při dosazení n a c vypočtených ze soustavy (6) pomocí hypergeometrického rozdělení:

$$normaHg = \|(o_1, o_2)\|_2$$

2) při dosazení n a c vypočtených ze soustavy (7) pomocí binomického rozdělení:

$$normaBi = \|(o_1, o_2)\|_2$$

3) při dosazení n a c vypočtených ze soustavy (8) pomocí Poissonova rozdělení:

$$normaPo = \|(o_1, o_2)\|_2$$

4) při dosazení n a c získaných z normy [3]

$$normaPlan = \|(o_1, o_2)\|_2 .$$

Pro porovnání vypočítaných výsledků s normou [3] je vytvořena tab. 3. Podle normy (v ní se u jednotlivých tabulek využívá označení pomocí přípustného a nepřípustného procenta vadných výrobků v dodávce P_1 a P_2) odpovídá pro $P_1 \in \langle 3{,}8; 4{,}5\rangle$ a $P_2 \in \langle 14; 18\rangle$ přejímací plán (65; 5). Tabulka ukazuje výsledky výpočtu přejímacích plánů (n; c), v krajních hodnotách těchto intervalů, při řešení pomocí hypergeometrického, binomického a Poissonova rozdělení podle vytvořeného programu (viz. přílohy Program2, popř. kapitola 3.3.4). Dále jsou v tabulce uvedeny normy vektorů odchylek pro jednotlivé způsoby výpočtu.

Tab. 3: Přejímací plán (65; 5) podle normy ČSN 01 0254

P_1	P_2	Po	Bi	Hg	normaHg	normaBi	normaPo	normaPlan
3,8	14	(67; 5)	(55; 4)	(54; 4)	0,0031	0,0064	0,0317	0,0235
3,8	18	(37; 3)	(35; 3)	(35; 3)	0,0103	0,0103	0,0241	0,0885
4,5	14	(75; 6)	(73; 6)	(63; 5)	0,0076	0,0140	0,0236	0,0213
4,5	18	(44; 4)	(43; 4)	(42; 4)	0,0141	0,0162	0,0232	0,0879

Jak je vidět z uvedené tabulky, nejpřesnější výsledky dává podle očekávání řešení pomocí hypergeometrické rozdělení, protože se z tohoto rozdělení vychází. Naopak přejímací plán (65; 5) uvedený v normě dává někdy značné odchylky od řešení podle hypergeometrického rozdělení. Tím je stanoven přejímací plán (n, c) odlišný od normy, který lépe aproximuje řešení rovnic (6).

Norma [3], se kterou je provedeno srovnávání výsledků, by ovšem měla být vytvořena při aproximaci Poissonovým rozdělením. Přejímací plán (65; 5), pro přípustné a nepřípustné procento vadných výrobků $P_1 \in \langle 3{,}8; 4{,}5 \rangle$ a $P_2 \in \langle 14; 18 \rangle$, tedy vznikl vhodnou aproximací přejímacích plánů v krajních hodnotách uvedených intervalů (67; 5), (37; 3), (75; 6) a (44; 4).

Z pozorování výsledků vytvořeného programu je možné usoudit, že hodnoty n a c přejímacího plánu (n; c), vypočítané podle řešení hypergeometrickým rozdělením, bývají zpravidla menší než při aproximaci Poissonovým rozdělením. Při aproximaci binomickým rozdělením leží získána čísla n a c někde mezi těmito čísly získanými podle ostatních dvou rozdělení.

Zajímavostí těchto tří způsobů řešení je, že pouze při řešení soustavy (6), tedy podle hypergeometrického rozdělení, je potřeba znát celkový počet výrobků v dodávce N. Jak moc ovlivňuje výpočet přejímacího plánu (n; c) celkový počet výrobků N v dodávce, ukazuje následující tabulka.

Tab. 4: Vliv celkového počtu výrobků N v dodávce (pro přejímací plán (17; 0) dle normy [3])

P_1	P_2	Po	Bi	Hg (N=100)	Hg (N=1000)	Hg (N=10000)
0,5	16	(14; 0)	(13; 0)	(21; 1)	(13; 0)	(13; 0)

Podle získaných výsledků lze usoudit, že vliv na výsledek mají především menší hodnoty N, kdežto při větších N (např. 1000, 10 000) jsou výsledky prakticky stejné.

Na rozdíl od přejímacích plánů v normě, počítá vytvořený program (viz. přílohy Program2, popř. kapitola 3.3.4) přejímací plány (n; c) přesně pro zadané vstupní parametry p_1, p_2, α a β.

3.3.4 Počítačový program pro výpočet přejímacích plánů

Pro výpočet optimálního přejímacího plánu (n; c) byl vytvořen počítačový program Program2 (viz. přílohy). Je naprogramován v prostředí Matlab R2010a metodou Switched board programming.

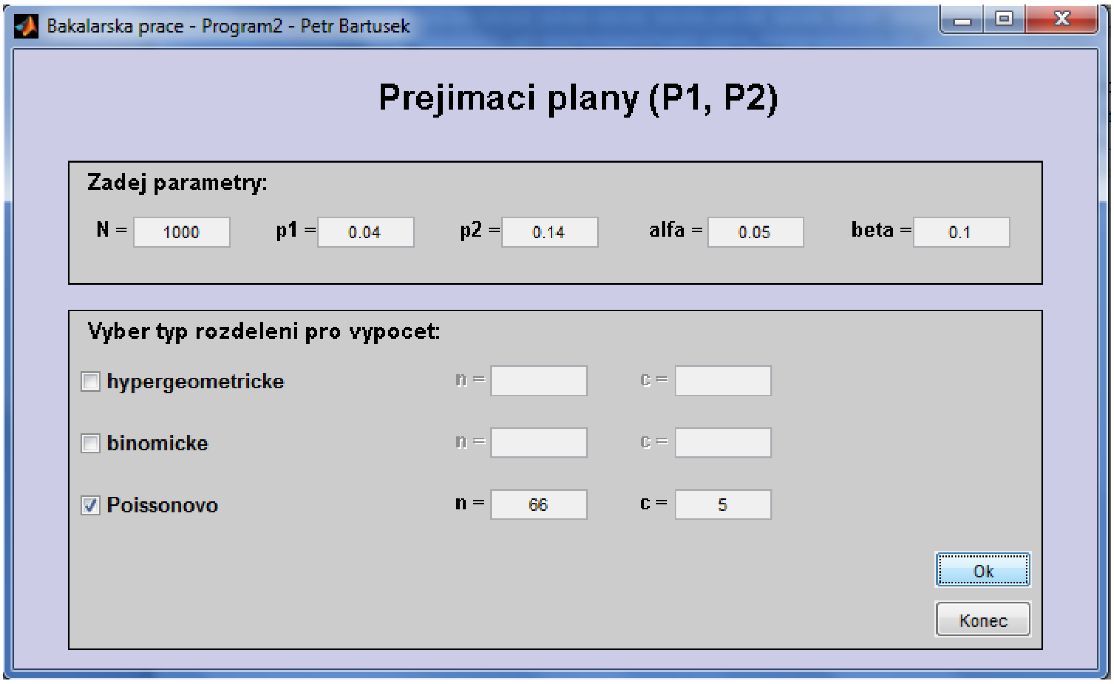
Při spuštění se uživateli objeví grafické okno. V jeho horní části zadává pět počátečních parametrů:

N – počet výrobků v celém souboru
p1 – přípustný podíl vadných výrobků v dodávce
p2 – nepřípustný podíl vadných výrobků v dodávce
alfa – riziko dodavatele (nastaveno na hodnotě $alfa = 0{,}05$)
beta – riziko odběratele (nastaveno na hodnotě $beta = 0{,}1$).

Hodnoty p1, p2, alfa, beta jsou reálná čísla z intervalu $(0; 1)$ a při jejich zadávání je potřeba dát pozor, že je nutné místo desetinné čárky psát tečku. Uživatel by měl volit čísla p1 a p2 z intervalu $p1 \in (0{,}001;\ 0{,}105)$, $p2 \in (0{,}0025;\ 0{,}35)$. Podle [7] je voleno riziko dodavatele $alfa = 0{,}05$ a riziko odběratele $beta = 0{,}1$.

V dolní části si uživatel vybírá podle jakého diskrétního rozdělení (hypergeometrické, binomické nebo Poissonovo) chce provést výpočet přejímacího plánu (n; c), s tím, že může zvolit více variant současně. Po stisknutí tlačítka Ok se objeví výsledné hodnoty. Čísla rozsahu výběru n a přejímacího čísla c jsou pak zobrazeny vpravo od vybraného diskrétního rozdělení. Program se ukončuje stisknutím tlačítka konec.

V programu jsou voleny hranice pro n a c podle kapitoly 3.3.3, tedy pro rozsah výběru platí $n \in \langle 0; 1000 \rangle$ a pro přejímací číslo $c \in \langle 0; 70 \rangle$. V případě potřeby širších intervalů, pro výpočet větších n, popř. c, by bylo potřeba provést změny ve zdrojovém kódu (viz. přílohy Program2). Při volbě aproximace Poissonovým rozdělením trvá výpočet podstatně kratší dobu. To je způsobeno tím, že je nejprve použita metoda bisekce pro přibližné určení c_p a poté je voleno $c \in \langle c_p - 5; c_p + 5 \rangle$. Naopak při volbě hypergeometrického nebo binomického rozdělení je nutné počítat s tím, že výpočet může trvat až několik desítek vteřin.



Obr. 18: Program2 – grafické prostředí pro výpočet přejímacího plánu (n; c) pro Poissonovo rozdělení

3.3.5 Počítačový program pro simulaci statistické přejímky srovnáváním

Nejprve se pro známé hodnoty N, p_1, p_2, α, β najde podle norem nebo podle vytvořeného programu Program2 přejímací plán (n, c). V simulaci je zkoumáno, kolik dodávek je zamítnuto ze 100 provedených kontrol, když podíl vadných výrobků v dodávce p je $p < p_1$, $p_1 < p < p_2$, $p_2 < p$.

Nechť z celkového počtu výrobků $N = 1000$ je K vadných. Výrobkům jsou přiřazena čísla 1 až N, kde prvních K je vadných. Ve vytvořeném programu (viz. přílohy program Simulace) se pro každou kontrolu vygeneruje n čísel (n je rozsah výběru z přejímacího plánu) pomocí generátoru náhodných čísel z intervalu (0; 1) zaokrouhlených na 3 desetinná místa a vynásobených 1000. Tím vznikne n náhodných celých čísel z intervalu (0;1000).

Je-li některé vygenerované číslo z intervalu $\langle 1; 50 \rangle$, byl při kontrole vybrán výrobek, který je zmetek. Nechť X je počet zjištěných zmetků při kontrole jedné dodávky. Rozhodnutí zda je taková dodávka přijata, či zamítnuta, probíhá podle kapitoly 3.3.2:

$X \leq c$ dodávka je přijata

$X > c$ dodávka je zamítnuta.

Výsledkem je zjištění počtu zamítnutých dodávek ze 100 kontrol a pravděpodobnost přijetí dodávky pro tři různé hodnoty podílu vadných výrobků v dodávce.

Nechť pro $N = 1000,\ p_1 = 0{,}04,\ p_2 = 0{,}15,\ \alpha = 0{,}05,\ \beta = 0{,}1$ je zjištěný

a) přejímací plán $(65; 5)$ podle normy [3]

Podíl vadných výrobků p	Počet zmetků K	Počet zamítnutí	Pravděpodobnost přijetí dodávky
0,03	30	1	0,99
0,09	90	56	0,44
0,16	160	96	0,04

Tab. 5: Provedení 100 kontrol simulace statistické přejímky srovnáváním pro tři různé hodnoty podílu vadných výrobků v dodávce při přejímacím plánu (65; 5)

b) přejímací plán $(51; 4)$ podle vytvořeného programu Program2 pro hypergeometrické rozdělení

Podíl vadných výrobků p	Počet zmetků K	Počet zamítnutí	Pravděpodobnost přijetí dodávky
0,03	30	2	0,98
0,09	90	49	0,51
0,16	160	93	0,07

Tab. 6: Provedení 100 kontrol simulace statistické přejímky srovnáváním pro tři různé hodnoty podílu vadných výrobků v dodávce při přejímacím plánu (51; 4)

Z výsledků simulací je vidět, že pro přejímací plán $(51; 4)$ je při menších hodnotách podílu vadných výrobků p pravděpodobnost přijetí dodávky menší než při přejímacím plánu (65; 5). Naopak při větších hodnotách podílu vadných výrobků p je pravděpodobnost přijetí dodávky menší u přejímacího plánu (65; 5). Hodnoty pravděpodobností přijetí dodávky, pro různé podíly vadných výrobků, odpovídají průběhu operativní charakteristiky pro daný přejímací plán (n, c).

Vytvořený program s názvem Simulace (viz. přílohy) nemá grafické rozhraní, je vytvořen pouze jako script v prostředí Matlab. Po spuštění dává jako výsledek počet zamítnutí dodávek ze 100 kontrol pro zvolené počáteční parametry N, p_1, p_2, α, β.

3.4 Další statistické přejímky srovnáváním

V následujících kapitolách jsou zmíněny další statistické přejímky srovnáváním, které se v praxi používají. Poznatky jsou získány zejména z [3],[4] a [7], kde je možné najít podrobnější informace.

3.4.1 Statistické přejímky pro ověření jakosti jednotlivého souboru

3.4.1.1 Bezopravné přejímací plány (P_1, P_2)

Přejímka dvojím výběrem

U statistické přejímky dvojím výběrem jsou na počátku známy následující údaje:

N – počet výrobků v celém souboru
p_1 – přípustný podíl vadných výrobků v dodávce
p_2 – nepřípustný podíl vadných výrobků v dodávce
α – riziko dodavatele
β – riziko odběratele,

kde p_1, p_2, α, β jsou voleny po domluvě dodavatele a odběratele. Z těchto údajů se vytvoří ideální přejímací plán (n_1, c_1, z_1, n_2, c_2).

n_1 – rozsah prvního výběru
c_1 – přípustný počet vadných kusů v prvním výběru
z_1 – nepřípustný počet vadných kusů v prvním výběru
n_2 – rozsah druhého výběru
c_2 – přípustný celkový počet vadných kusů v obou výběrech dohromady

Při statistické přejímce dvojím výběrem se nejprve vybírá z celého souboru N výrobků náhodně n_1 výrobků. Tyto výrobky se překontrolují a zjistí se počet vadných kusů. Pokud je toto číslo menší nebo rovno než přípustný počet vadných kusů v prvním výběru, pak je dodávka ihned přijata. Pokud je počet vadných kusů větší nebo roven než nepřípustný počet vadných kusů v prvním výběru, je dodávka naopak ihned zamítnuta. Jestliže je počet vadných mezi těmito dvěma hodnotami, pokračuje se druhým náhodným výběrem n_2 výrobků. Platí, že $n_1 = n_2$.

Druhý náhodný výběr probíhá z těch výrobků, které zbyly v dodávce po prvním výběru. Druhý výběr projde kontrolou a zjistí se počet vadných kusů. Pokud je součet nevyhovujících výrobků v prvním a druhém výběru dohromady menší nebo roven než přípustný celkový počet vadných c_2, je dodávka přijata, v opačném případě je zamítnuta.

Nechť X_1 je počet vadných kusů v prvém výběru a X_2 počet vadných kusů v druhém výběr. Postup statistické přejímky dvojím výběrem je následující

$X_1 \leq c_1$ soubor se přijímá
$X_1 \geq z_1$ soubor se zamítá
$c_1 < X_1 < z_1$ provede se druhý náhodný výběr

$X_1 + X_2 \leq c_2$ soubor se přijímá
$X_1 + X_2 > c_2$ soubor se zamítá.

Přejímka několikerým výběrem

U statistické přejímky několikerým výběrem jsou na počátku známy následující údaje:

N – počet výrobků v celém souboru
p_1 – přípustný podíl vadných výrobků v dodávce
p_2 – nepřípustný podíl vadných výrobků v dodávce
α – riziko dodavatele
β – riziko odběratele
k - předepsaný maximální počet výběrů,

kde p_1, p_2, α, β jsou voleny po domluvě dodavatele a odběratele. Z těchto údajů se vytvoří ideální přejímací plán ($n_1, \ldots, n_k, c_1, \ldots, c_k, z_1, \ldots, z_{k-1}$).

n_i – rozsah i-tého výběru

c_i – přípustný počet vadných kusů v prvním až v i-tém výběru

z_i – nepřípustný počet vadných kusů v prvním až v i-tém výběru

Postupuje se analogicky jako o přejímky dvojím výběrem. Pro náhodné výběry platí $n_1 = n_2 = ... = n_k$. Nechť X_i je počet vadných kusů v i-tém náhodném výběru a tedy po žádném z předchozích výběrů nedošlo k rozhodnutí o přijetí nebo zamítnutí. Postupuje se dále následovně:

$X_1 + \ldots + X_i \leq c_i$	soubor se přijímá
$X_1 + \ldots + X_i \geq z_i$	soubor se zamítá
$c_i < X_1 + \ldots + X_i < z_i$	provede se (i+1)-vý náhodný výběr.

Jestliže k rozhodnutí o přijetí nedojde až do k-tého výběru, pak po náhodném výběru n_k se už rozhodne definitivně:

$X_1 + \ldots + X_k \leq c_k$	soubor se přijímá
$X_1 + \ldots + X_k > c_k$	soubor se zamítá.

Přejímka sekvenční

U statistické přejímky sekvenční jsou na počátku známy následující údaje:

N – počet výrobků v celém souboru

p_1 – přípustný podíl vadných výrobků v dodávce

p_2 – nepřípustný podíl vadných výrobků v dodávce

α – riziko dodavatele

β – riziko odběratele,

kde p_1, p_2, α, β jsou voleny po domluvě dodavatele a odběratele. Vytvoří se přejímací plán, v němž A, a_1, a_2, d jsou konstanty závislé na p_1, p_2, α, β (viz. [7] str. 38).

k_i – celkový počet vadných kusů po i-tém výběru

$a_1 + id$ - hranice pro přijetí při i-tém výběru

$a_2 + id$ - hranice pro zamítnutí při i-tém výběru

Při sekvenční přejímce se v každém výběru vybere náhodně vždy jen jeden prvek. Je-li výrobek v pořádku, přiřadí se mu hodnota 0, je-li vadný, přiřadí se mu hodnota 1. Celkový počet vadných kusů po i-tém výběru dostaneme tak, že

$$k_i = \sum_{j=1}^{i} x_j$$

Po každém překontrolování výrobku se k_i porovná s hranící pro přijetí a pro zamítnutí. Jestliže je hodnota mezi těmito dvěma hranicemi, pokračuje se náhodným výběrem dalšího výrobku.

Postup při i-tém výběru lze tedy zapsat následovně:

$k_i \leq a_1 + id$	soubor se přijímá
$k_i \geq a_2 + id$	soubor se zamítá
$a_1 + id < k_i < a_2 + id$	pokračuje se výběrem dalšího kusu ke kontrole.

Hranice pro přijetí a zamítnutí lze graficky interpretovat jako polopřímky s definičním oborem $\langle 0, \infty)$, procházející I. a IV. kvadrantem souřadné soustavy.

$y = a_1 + dx$	hranice pro přijetí
$y = a_2 + dx$	hranice pro zamítnutí

Pokud by dlouho nedošlo k rozhodnutí o přijetí nebo zamítnutí, postupuje se tak, že se najde přejímací plán pro jeden výběr (n, c). Pokud po n-tém výběru při sekvenční přejímce nedojde k rozhodnutí, rozhodne se dle přejímacího plánu pro jeden výběr. Nedělá se však nový náhodný výběr n výrobků, ale využije se těch n náhodně vybraných kusů z již uskutečněné sekvenční přejímky.

$k_n \leq c$ soubor se přijímá
$k_n > c$ soubor se zamítá

3.4.1.2 Opravné přejímací plány (P_2)

Opravné přejímací plány zajišťují splnění podmínky $L(p_2) = \beta$ a minimalizují střední počet kontrolovaných výrobků při průměrném podílu vadných v dodávce.

U opravných přejímacích plánů jsou na počátku známy následující údaje:

N – počet výrobků v celém souboru
p_2 – nepřípustný podíl vadných výrobků v dodávce
$\bar{p}$ – průměrný podíl vadných výrobků v dodávce
β – riziko odběratele,

kde p_2, β jsou voleny po domluvě dodavatele a odběratele, $\bar{p}$ je zjištěn z dříve uskutečněných přejímek. Z těchto údajů se vytvoří ideální přejímací plán (n, c) u opravných přejímacích plánů jedním výběrem

n – rozsah výběru
c – akceptační číslo (přípustný počet vadných kusů ve výběru),

popř. (n_1, c_1, n_2, c_2) u opravných přejímacích plánů dvojím výběrem.

n_1 – rozsah prvního výběru
n_2 – rozsah druhého výběru
c_1 – přípustný počet vadných kusů v prvním výběru
c_2 – přípustný počet vadných kusů v prvním a druhém výběru dohromady

Rozdíl oproti bezopravným přejímacím plánům je v tom, co se děje s vadnými kusy. Kontrolou zjištěné nevyhovující výrobky při náhodném výběru musí dodavatel opravit nebo nahradit novými. Pokud je jejich počet menší nebo roven přípustnému počtu vadných kusů ve výběru c, pak je dodávka s nahrazenými výrobky přijata. Je-li počet nevyhovujících výrobků při kontrole větší než c, pak je zkontrolován celý zbytek souboru a vadné kusy jsou dodavatelem opět opraveny nebo nahrazeny novými. Až poté je dodávka přijata.

V přejímacím plánu pro přejímku dvojím výběrem se neuvádí číslo z_1 vyjadřující nepřípustný počet vadných kusů v prvním výběru. To je z důvodu, že pro toto číslo pokaždé platí

$$z_1 = c_2 + 1$$

Při hledání přejímacího plánu v tabulkách jsou tyto tabulky rozděleny pro různé hodnoty průměrného procenta vadných výrobků P_s a nepřípustného procenta vadných výrobků P_2. Tyto hodnoty se dostanou lehce ze vztahů:

$$P_s = \bar{p} \cdot 100$$
$$P_2 = p_2 \cdot 100$$

3.4.2 Statistické přejímky pro ověření jakosti souborů tvořících sérii

3.4.2.1 Opravné přejímací plány (P_L)

Na rozdíl oproti opravným přejímacím plánům (P_2) pro kontrolu jakosti jednotlivého souboru se zde neuvažuje podmínka $L(p_2) = \beta$. Pro zachování záruky odběratele se zavádí *průměrná výstupní jakost* AOQ (average outgoing quality). Jestliže je před kontrolou podíl vadných výrobků v dodávce p, pak funkce AOQ(p) udává střední podíl vadných kusů po kontrole. Maximum této funkce se nazývá *mez průměrné výstupní jakosti* a značí se p_L, popř. AOQL (average outgoing quality limit). Z toho vyplývá, že průměrná výstupní jakost nebude horší než mez průměrné výstupní jakosti.

Opravné přejímací plány (P_L) zajišťují splnění podmínky

$$\max_{0<p<1} AOQ\,(p) = p_L$$

a minimalizují střední počet kontrolovaných výrobků při průměrném podílu vadných v dodávce I_s.

U opravných přejímacích plánů (P_L) jsou na počátku známy následující údaje:

N – počet výrobků v celém souboru
p_L – mez průměrné výstupní jakosti
$\bar{p}$ – průměrný podíl vadných výrobků v dodávce,

kde p_L je volena po domluvě dodavatele a odběratele, $\bar{p}$ je zjištěn z dříve uskutečněných přejímek. Z těchto údajů se vytvoří ideální přejímací plán (n, c) u opravných přejímacích plánů jedním výběrem

n – rozsah výběru
c – akceptační číslo (přípustný počet vadných kusů ve výběru),

popř. (n_1, c_1, n_2, c_2) u opravných přejímacích plánů dvojím výběrem.

n_1 – rozsah prvního výběru
n_2 – rozsah druhého výběru
c_1 – přípustný počet vadných kusů v prvním výběru
c_2 – přípustný počet vadných kusů v prvním a v druhém výběru dohromady

Průběh přejímací kontroly, i to co se děje se zjištěnými vadnými kusy, je obdobný jako u opravných přejímacích plánů (P_2).

V přejímacím plánu pro přejímku dvojím výběrem se opět neuvádí číslo z_1 vyjadřující nepřípustný počet vadných kusů v prvním výběru. To je z důvodu, že pro toto číslo pokaždé platí

$$z_1 = c_2 + 1$$

Při hledání přejímacího plánu v tabulkách jsou tyto tabulky rozděleny pro různé hodnoty průměrného procenta vadných výrobků P_s a nejhoršího průměrného výstupního procenta vadných P_L. Tyto hodnoty se dostanou podle vztahů:

$$P_s = \bar{p} \cdot 100$$
$$P_L = p_L \cdot 100$$

3.4.2.2 Bezopravné přejímací plány AQL

Označení dle ISO 2859:

Ac - přejímací číslo (acceptance number)
Re - zamítací číslo (rejection number)
N - rozsah dávky
D - počet neshodných jednotek v dávce
n - rozsah výběru
d - počet neshodných jednotek zjištěný ve výběru z dávky
p - průměr procesu
p_x - úroveň jakosti, pro kterou je pravděpodobnost přijetí rovna x, kde $x \in \langle 0,1 \rangle$
AQL - přípustná mez jakosti (acceptance quality limit) – úroveň jakosti, která je nejhorším přijatelným průměrem procesu
CRQ - jakost odpovídající riziku odběratele (consumer´s risk quality)
AOQ - průměrná výstupní jakost
AOQL - nejhorší průměrná výstupní jakost

Přejímací plány AQL zajišťují, že pokud není průměrné procento vadných v souboru tvořícím sérii vyšší než AQL, pak odběratel přijme dodávku podle přejímacího plánu s danou hodnotou AQL s vysokou pravděpodobností.

Kontrolní úrovně ukazují závislost mezi rozsahem výběru n a rozsahem dodávky N. Kontrolní úrovně se rozlišují tři základní - I, II, III a čtyři speciální - S_1, S_2, S_3, S_4. Za běžných okolností se používá úroveň II.

Na dané kontrolní úrovni platí určitá přechodová pravidla pro přestup na jiný typ kontrol. *Typy kontrol* jsou tři – normální, zpřísněná, zmírněná. Za běžných podmínek se používá normální kontrola. Zpřísněná kontrola je zavedena pokud průměr procesu je horší než AQL, zmírněná kontrola je zavedena pokud průměr procesu je lepší než AQL. Dle [4] platí:

Normální kontrola

- při zahájení kontroly se vždy začíná normální kontrolou
- jestliže nejsou přijaty 2 z 5 po sobě jdoucích souborů, přechází se na zpřísněnou kontrolu
- jestliže je stabilizovaná výroba, počet bodů pro přechod je větší nebo roven 30 a pověřený orgán přechod schválil, přechází se na zmírněnou kontrolu

Zpřísněná kontrola

- jestliže není přijato 5 dávek během zpřísněné kontroly, dochází k přerušení kontroly, dokud dodavatel nezajistí zlepšení jakosti výrobků
- jestliže je přijato 5 po sobě jdoucích souborů, přechází se zpět na normální kontrolu

Zmírněná kontrola

- jestliže jeden soubor není přijat, výroba je nepravidelná, popř. z jiných závažných důvodu, přechází se zpět na normální kontrolu

U bezopravných přejímacích plánů AQL jsou na počátku známy následující údaje:

N – počet výrobků v celém souboru
AQL – přípustná mez jakosti
kontrolní úroveň.

Z těchto údajů se vytvoří ideální přejímací plán, který je pro přejímku jedním výběrem dán kombinací čísel

n – rozsah výběru
Ac – přípustný počet vadných kusů ve výběru
Re – nepřípustný počet vadných kusů ve výběru,

pro přejímku dvojím výběrem

n_1 - rozsah prvního výběru
n_2 - rozsah druhého výběru
Ac_1 - přípustný počet vadných kusů v prvním výběru
Re_1 - nepřípustný počet vadných kusů v prvním výběru
Ac_2 - přípustný počet vadných kusů v obou výběrech dohromady
Re2 - nepřípustný počet vadných kusů v obou výběrech dohromady,

popř. pro přejímku několikerým výběrem

n_i - rozsah i-tého výběru
Ac_i - přípustný celkový počet vadných kusů v prvním až i-tém výběru
Re_i - nepřípustný celkový počet vadných kusů v prvním až i-tém výběru.

Při hledání přejímacího plánu jedním výběrem podle tabulek se postupuje následovně. Dle rozsahu dávky N a kontrolní úrovně se najde odpovídající písmeno A, B, C,…,R. Poté se v další tabulce dané typem kontroly najdou pro toto písmeno a příslušnou přípustnou mez jakosti AQL přejímací číslo Ac, zamítací číslo Re a rozsah výběru n. Jestliže je v tabulce místo Ac a Re šipka, pak tato šipka směřuje k takovému přejímací a zamítacímu číslu, které se mají použít. Analogicky se hledají přejímací plány pro přejímky dvojím a několikerým výběrem.

Rozhodovací kritéria, zda dodávku přijmout či zamítnout, jsou pro soubory tvořící sérii stejná jako u izolovaných dávek.

4. Závěr

Cílem bakalářské práce je popis statistických přejímek srovnáváním a odvození konstrukce přejímacího plánu pro přejímky s jedním výběrem.

Na začátku bakalářské práce jsou vysvětleny matematické pojmy, které se dále využívají. Druhá kapitola se zabývá popisem tří diskrétních rozdělení pravděpodobností – hypergeometrického, binomického a Poissonova, která jsou následně potřeba při konstrukcích přejímacích plánů. U všech rozdělení je uvedena pravděpodobnostní funkce, distribuční funkce a vypočítány číselné charakteristiky. Pro lepší názornost těchto pojmů byl k této kapitole vytvořen v prostředí Matlab program (viz. přílohy Program1), ve kterém si lze navíc prohlédnout ukázky konvergencí jednotlivých rozdělení pravděpodobností. Aproximacím jednotlivých diskrétních rozdělení pravděpodobností je věnován konec druhé kapitoly, kde je odvozeno, že hypergeometrické rozdělení lze aproximovat binomickým, je-li splněna podmínka

$$\frac{n}{N} < 0{,}1\,.$$

Binomické rozdělení lze aproximovat Poissonovým, jsou-li splněny podmínky

$$n > 30, \;\; \theta < 0{,}1\,.$$

Při splnění všech tří podmínek je možné aproximovat hypergeometrické rozdělení Poissonovým.

Třetí kapitola začíná popisem statistických přejímek. Ty se využívají při výběrové kontrole jakosti výrobků. Po vysvětlení základní pojmů je uvedeno několik kritérií, podle kterých lze statistické přejímky dělit. Poté už je pozornost věnována hlavnímu cíli bakalářské práce, tedy popisu statistických přejímek srovnáváním s jedním výběrem. Pro tento typ přejímky jsou odvozeny tři způsoby výpočtu přejímacího plánu (n, c), podle zvoleného diskrétního rozdělení pravděpodobností. Pro výpočet přejímacího plánu (n, c) je opět vytvořen v prostředí Matlab program (viz. přílohy Program2). Jelikož přesné řešení vznikne vyřešením soustavy rovnic s hypergeometrickým rozdělením

$$F_{Hg}(c; N, Np_1, n) = \sum_{i=0}^{c} \frac{\binom{Np_1}{i}\binom{N-Np_1}{n-i}}{\binom{N}{n}} = 1 - \alpha$$

$$F_{Hg}(c; N, Np_2, n) = \sum_{i=0}^{c} \frac{\binom{Np_2}{i}\binom{N-Np_2}{n-i}}{\binom{N}{n}} = \beta\,,$$

jsou s tímto řešením srovnány i další řešení pomocí binomického rozdělení a Poissonova rozdělení. Vypočítané přejímací plány z vytvořeného programu jsou porovnány s přejímacími plány uvedenými v normě [3]. Je zjištěno, že v uvedené normě je jeden přejímací plán pro určité hranice přípustného a nepřípustného podílu vadných výrobků v dodávce p_1 a p_2, kdežto vytvořený program Program2 je schopný určit přesný přejímací plán pro všechny dvojice p_1 a p_2. Ke třetí kapitole je navíc vytvořen simulační program (viz. přílohy Simulace), který simuluje výběrovou kontrolu pro statistické přejímky srovnáváním. V závěru třetí kapitoly jsou uvedeny další statistické přejímky srovnáváním, které se v praxi používají.

Všechny tři vytvořené programy jsou naprogramovány v prostředí Matlab a jsou přiloženy na CD.

5. Zdroje

[1] Anděl, J.: Matematická statistika. STNL/ALFA, Praha, 1978.
[2] ČERMÁK, L. – HLAVIČKA, R.: Numerické metody. Akademické nakladatelství CERM, Brno, 2008.
[3] Československá státní norma ČSN 01 0254. Statistická přejímka srovnáváním. Vydavatelství úřadu pro normalizaci a měření, Praha, 1983.

[4] Česká technická norma ČSN ISO 2859-1 01 0261. Statistické přejímky srovnáváním - Část 1: přejímací plány AQL pro kontrolu každé dávky v sérii. Český normalizační institut, Praha, 2000.

[5] HÁTLE, J. – LIKEŠ, J.: Základy počtu pravděpodobnosti a matematické statistiky. SNTL, Praha, 1974.

[6] KARPÍŠEK, Z.: Matematika IV Statistika a pravděpodobnost. Akademické nakladatelství CERM, Brno, 2007.

[7] KLŮFA, J.: Opravné statistické přejímky. Ekopress, Praha, 2009.

[8] LIKEŠ, J. – MACHEK, J.: Matematická statistika. SNTL, Praha, 1983.

[9] ZAPLATÍLEK, K. – DOŇAR, B.: MATLAB – tvorba uživatelských aplikací. Ben – technická literatura, Praha, 2004.

[10] http://en.wikipedia.org/wiki/Hypergeometric_distribution

[11] http://en.wikipedia.org/wiki/Binomial_distribution

[12] http://en.wikipedia.org/wiki/Poisson_distribution

[13] http://en.wikipedia.org/wiki/Moment-generating_function

6. Přílohy

Všechny tři programy (Program1, Program2, Simulace) jsou přiloženy na CD. Program1 se spustí ze souboru Program1.m a , Program2 ze souboru Program2.m a program Simulace se spustí ze souboru Simulace.m.