



1. domácí úloha

Výkonnost a spolehlivost prog. systémů

Ondřej Drtina
A20N0077P
drtinao@students.zcu.cz

Obsah

1	Zadání	2
1.1	Celé znění úlohy	2
2	Analýza problému	3
2.1	Základní informace	3
2.2	Distribuční funkce	4
2.3	Inverzní fce k distribuční funkci	4
2.4	Výpočet střední hodnoty	4
2.5	Výpočet rozptylu	5
3	Popis implementace	6
3.1	Starter.java	6
3.2	StatisticsCalc.java	6
3.3	GeometricDistribution.java	6
3.4	Použitý programovací jazyk, vývojové prostředí	6
4	Uživatelská příručka	7
4.1	Demo režim	7
4.2	Klasický režim	7
4.3	Klasický režim s volbou velikosti histogramu / počtu hvězd . .	7
4.4	Spuštění programu	7
5	Závěr	8

1 Zadání

V rámci první úlohy bylo cílem vytvořit generátor, jež bude poskytovat čísla náležící do zadaného pravděpodobnostního rozdělení. V mém případě je požadována implementace vztahující se k geometrickému rozdělení. Na implementaci byly kladeny následující základní požadavky:

- musí být využita vhodná metoda (inverzní transformace, zamítací metoda atp.)
- program bude možno spustit příkazem, jež odpovídá formátu: `program (počet generovaných čísel) (parametr rozdělení1) [parametr rozdělení2] [...]`
- pokud bude program spuštěn bez parametrů, budou provedeny dvě série výpočtů s předem připravenými demo parametry
- program bude dodán ve formě ZIP archivu, který bude po rozbalení možno okamžitě spustit na Windows / Linux
- program bude dodán se skriptem "run.bat", jež zajistí spuštění programu a předání parametrů

1.1 Celé znění úlohy

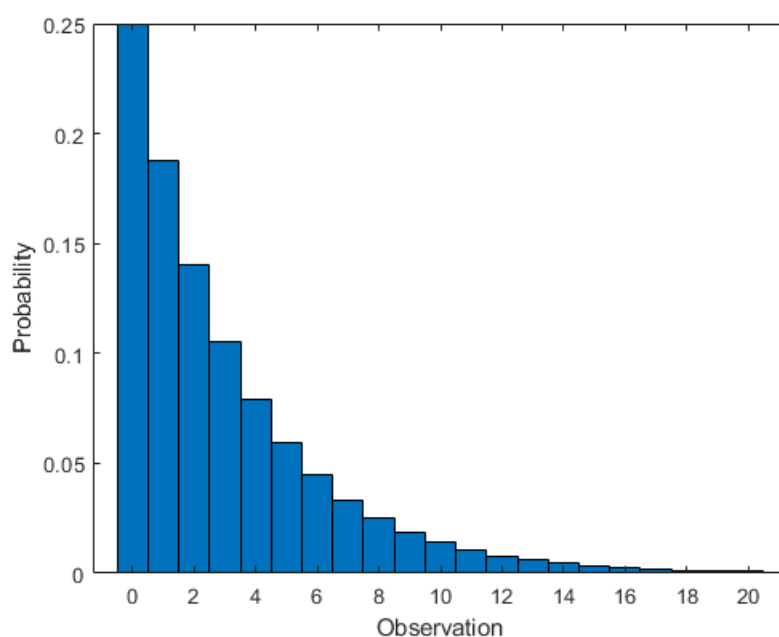
Celé zadání úlohy je možno nalézt online zde.

2 Analýza problému

Po přečtení zadání úlohy bylo jasné, že její splnění bude vyžadovat znalost vlastností geometrického rozdělení. Zopakoval jsem si tedy vlastnosti daného rozdělení, jehož popis je předmětem následujících kapitol.

2.1 Základní informace

Jedná se o jedno z diskrétních rozdělení, jež vyjadřuje počet neúspěšných pokusů, které nastanou před prvním úspěšným pokusem. Předpokládáme, že jednotlivé pokusy jsou na sobě nezávislé. Standardní geometrické rozdělení je definováno pro celočíselné hodnoty od nuly do nekonečna. Vizuálně lze graf ohodnotit jako klesající, tedy je zřejmé, že s každým dalším pokusem se pravděpodobnost snižuje (viz 2.1). Vodorovná osa vždy znázorňuje pořadí pokusu, svislá pak udává očekávanou procentuální úspěšnost daného pokusu.



Obrázek 2.1: Ukázkový graf geometrického rozdělení

2.2 Distribuční funkce

Distribuční funkce geometrického rozdělení je definována jako $F(k) = 1 - (1 - p)^{k+1}$, kde:

- p - pravděpodobnost úspěchu [%]
- k - pořadí pokusu

2.3 Inverzní fce k distribuční funkci

Úpravami lze z distribuční funkce získat inverzní funkci, která je nutná pro generování náhodných čísel spadajících do geometrického rozdělení. Inverzní funkce je definována jako $F^{-1} = \frac{\log(1-u)}{\log(1-p)} - 1$, kde:

- u - náhodné desetinné číslo $0 < u < 1$
- p - pravděpodobnost úspěchu [%]

Technika využití inverzní funkce pro generování náhodných čísel z rozdělení se označuje jako tzv. "inverzní transformace". Zamítací metodu v tomto případě není možno využít, jelikož distribuční funkce nemá jasně dané meze, není omezena zprava.

2.4 Výpočet střední hodnoty

Ideální střední hodnotu lze získat vzorcem $\frac{1-p}{p}$, kde:

- p - pravděpodobnost úspěchu [%]

V kódu musí být přítomen i obecný vzorec pro výpočet, který lze využít u všech rozdělení (viz 2.2):

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

Obrázek 2.2: Vzorec pro výpočet střední hodnoty

, kde:

- x_i - konkrétní prvek spadající do daného rozdělení
- n - počet všech vygenerovaných prvků

2.5 Výpočet rozptylu

Ideální rozptyl geometrického rozdělení lze získat vzorcem $\frac{1-p}{p*p}$, kde:

- p - pravděpodobnost úspěchu [%]

V kódu je bude pro kontrolu rozptylu využít i obecný vzorec (viz 2.3):

$$s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Obrázek 2.3: Vzorec pro výpočet rozptylu

, kde:

- x_i - konkrétní prvek spadající do daného rozdělení
- \bar{x} - střední hodnota
- n - počet všech vygenerovaných prvků

3 Popis implementace

Předmětem kapitoly je základní popis tříd v přítomných v programu. Metody zde nejsou popsány, jelikož zdrojový kód je důkladně okomentován.

3.1 Starter.java

Vstupní bod aplikace, obsahuje metodu `main`. Při spuštění programu je přítomnými metodami zkontrolován počet parametrů programu. Ke spuštění výpočtů dojde pouze tehdy, pokud nejsou žádné zadány (demo režim) či pokud jsou zadány validní parametry. Validní parametry viz kapitola 4.

Při validním vstupu dojde k volání metod z ostatních tříd, jež zajistí výpočty. Následně je výsledek metodou uvedené třídy vypsán uživateli do konzole.

3.2 StatisticsCalc.java

Třída zajišťující výpočet střední hodnoty a rozptylu pomocí histogramu. Jedná se o hodnoty `E_vypocet` a `D_vypocet`.

3.3 GeometricDistribution.java

Zajišťuje výpočet střední hodnoty a rozptylu pomocí vzorců, tedy zodpovídá za výpočet `E_teorie` a `D_teorie`.

3.4 Použitý programovací jazyk, vývojové prostředí

Program jsem vytvořil v programovacím jazyce Java a vývoj probíhal v rámci prostředí IntelliJ IDEA. Zmíněný jazyk jsem zvolil hlavně kvůli jeho přenositelnosti.

4 Uživatelská příručka

Program je možné spustit ve 3 režimech.

4.1 Demo režim

Demo režim se spustí, pokud uživatel nezadá žádné argumenty a provede výpočet na dvou předem definovaných množinách hodnot.

Spuštění v tomto režimu: `run.bat`.

4.2 Klasický režim

V tomto režimu jsou od uživatele očekávány hodnoty `p` a `x` (v tomto pořadí). Program následně provede výpočet se zadanými parametry.

Spuštění v tomto režimu např.: `run.bat 1000000 0.4`.

4.3 Klasický režim s volbou velikosti histogramu / počtu hvězd

V tomto režimu jsou od uživatele očekávány hodnoty `p`, `x` a `velikost_histogramu;max_pocet_hvezd` (v tomto pořadí).

Poslední hodnota udává počet nejvíce relevantních položek histogramu, jež se vypíše do terminálu a počet hvězd, které budou přiřazeny nejvíce frekvencovanému číslu. Počet hvězd ostatních položek je pak samozřejmě poměrný k dané hodnotě.

Spuštění v tomto režimu např.: `run.bat 1000000 0.4 10;100`.

4.4 Spuštění programu

Je dodán soubor `run.bat`, který umožňuje spuštění programu na PC s Linuxem i Windows - parametry viz předchozí kapitoly.

5 Závěr

Program byl otestován na OS Windows 10 Pro + Linux Mint. Funguje dle očekávání a generovaná čísla spadají do geometrického rozdělení.