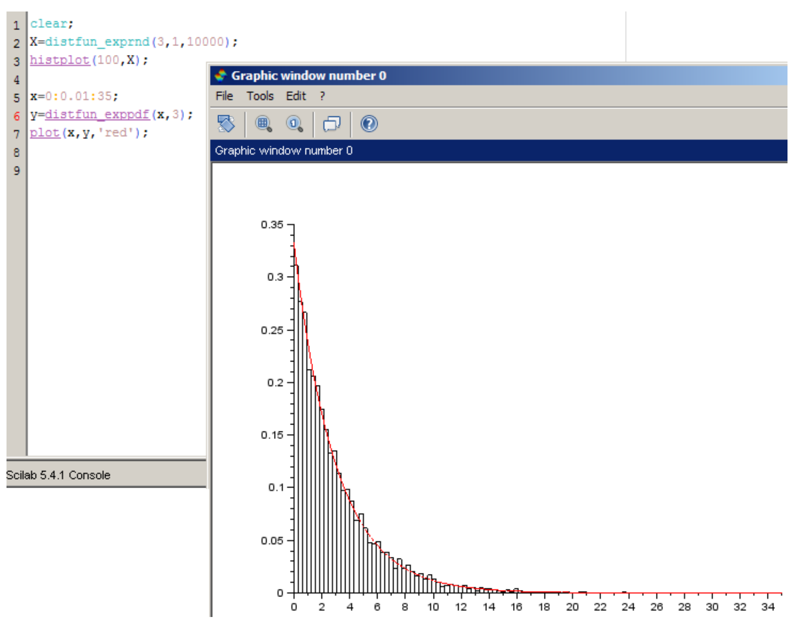
Tři hlavní věty pravděpodobnosti

15. května 2015

První příklad

Představme si, že máme atomy typu A, které se samovolným radioaktivním rozpadem rozpadají na atomy typu B. Průměrná doba rozpadu je 3 hodiny. Víme, že takový rozpad má exponenciální rozdělení.

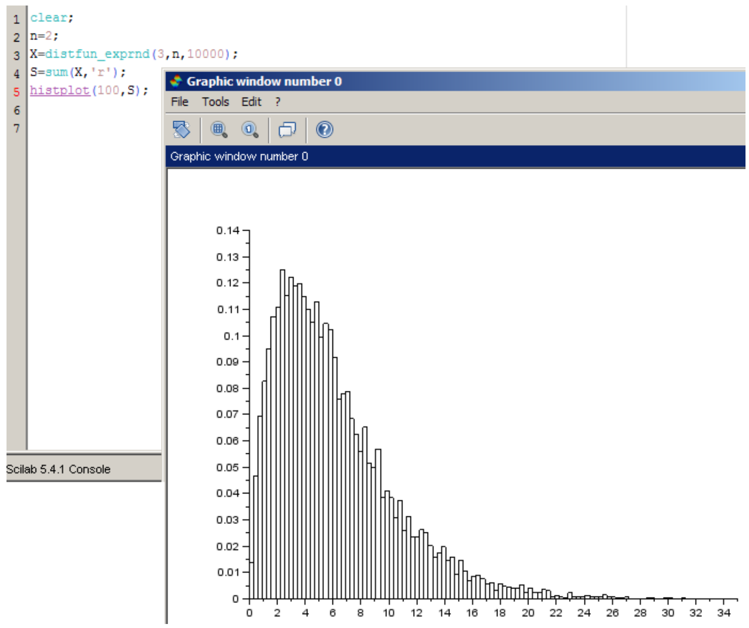
Provedeme ve Scilabu simulaci pro 10 000 jader, vykreslíme histogram a porovnáme s danou hustotou pravděpodobnosti $f(x) = \frac{1}{D} \cdot e^{-\frac{x}{D}}$, kde $D = 3$:



Až dosud pracujeme se známým rozdělením, takže např. spočítat, za jak dlouho se rozpadne 90% jader, není problém.

Nyní předchozí příklad upravme. Jádro A se bude rozpadat na jádro B a to na jádro C. Pro jednoduchost mějme oba rozpady exponenciální se střední dobou rozpadu 3 hodiny. Vykresleme nyní histogram pro náhodnou veličinu, za jak dlouho nám vznikne jádro C. Tato doba je dána součtem dob obou rozpadů - exponenciálních náhodných veličin.

V simulaci tedy vygenerujeme dva řádky náhodných čísel s příslušným exponenciálním rozdělením a ty sečteme. Pro tento součet pak vykreslíme histogram:



Vidíme, že toto rozdělení již rozhodně není exponenciální, ani nepřipomíná žádné jiné standardní rozdělení. Určit, za jak dlouho nám vznikne 90% jader C, je již mnohem náročnější problém.

V tomto případě bychom ještě uměli výslednou hustotu pravděpodobnosti $g(x)$ určit dle vzorce:

$$g(x) = \int_{t=0}^{x} f(t) \cdot f(x-t)\, dt,$$

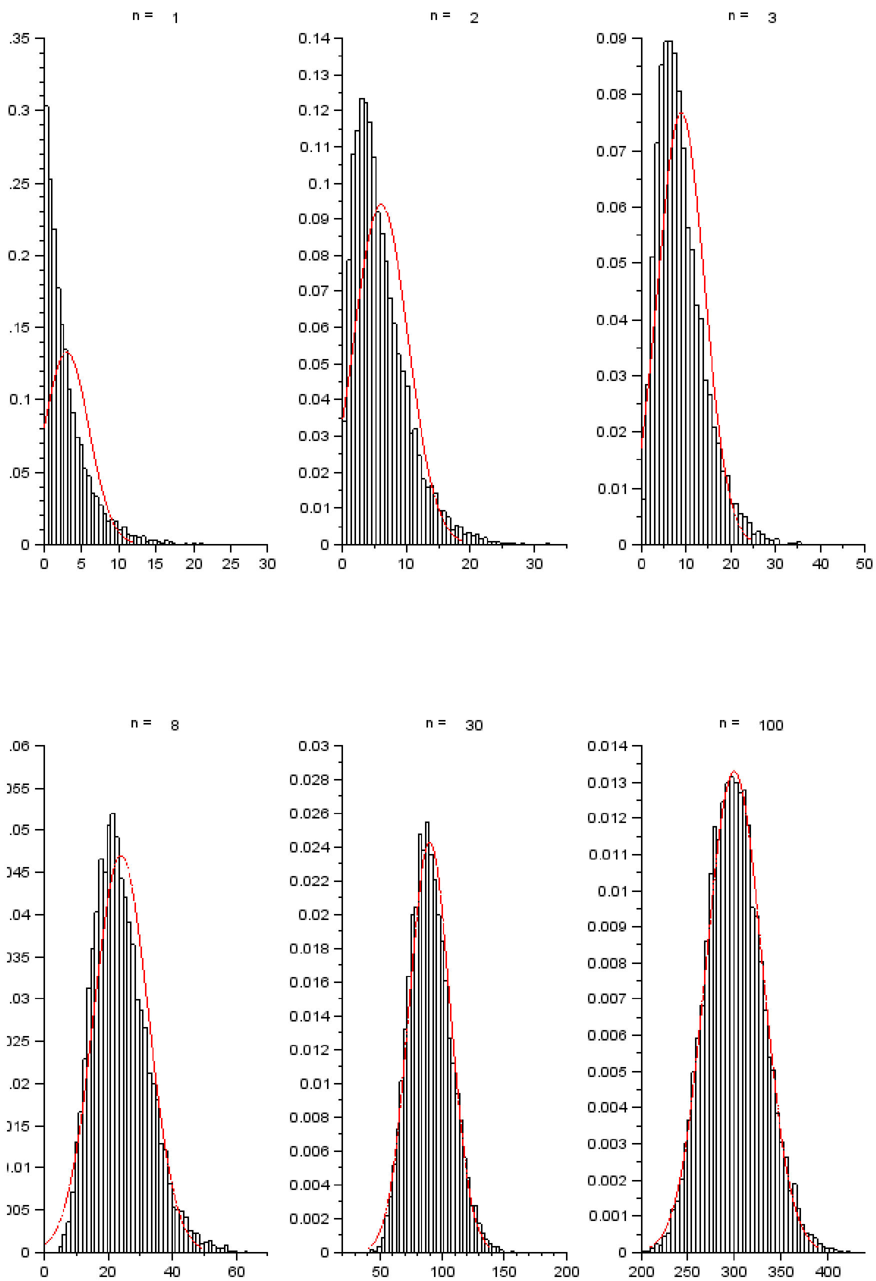
kde $f(t)$ je v našem případě hustota exponenciálního rozdělení.

Pokud bychom ale sčítali více náhodných veličin, tak s každou další by nám ve výpočtu přibýval jeden integrál a výpočet by se stal prakticky nezvládnutelným.

Vykresleme nyní histogramy, pokud by v rozpadové řadě bylo n rozpadů.

Červeně vykreslujeme Gaussovo rozdělení se stejnou střední hodnotou a rozptylem, jako má zkoumané součtové rozdělení. Tuto střední hodnotu a rozptyl spočteme velmi snadno, protože víme, že při sčítání náhodných veličin se střední hodnoty i rozptyly také sčítají. V našem případě budou tedy střední hodnota a rozptyl rovny:

$$\begin{aligned} \mu &= n \cdot D, \\ \rho^2 &= n \cdot D^2. \end{aligned}$$



Co pozorujeme? Začali jsme s exponenciálním rozdělením. Při sčítání byl tvar součtového rozdělení poměrně složitý, velmi rychle se však blížil normálnímu rozdělení. Pro $n = 8$ je součtové rozdělení ještě trochu nahnuté doleva, pro třicet je už víceméně dobré, pro sto už sedí výborně. Dokonale by samozřejmě sedělo až pro $n = \infty$.

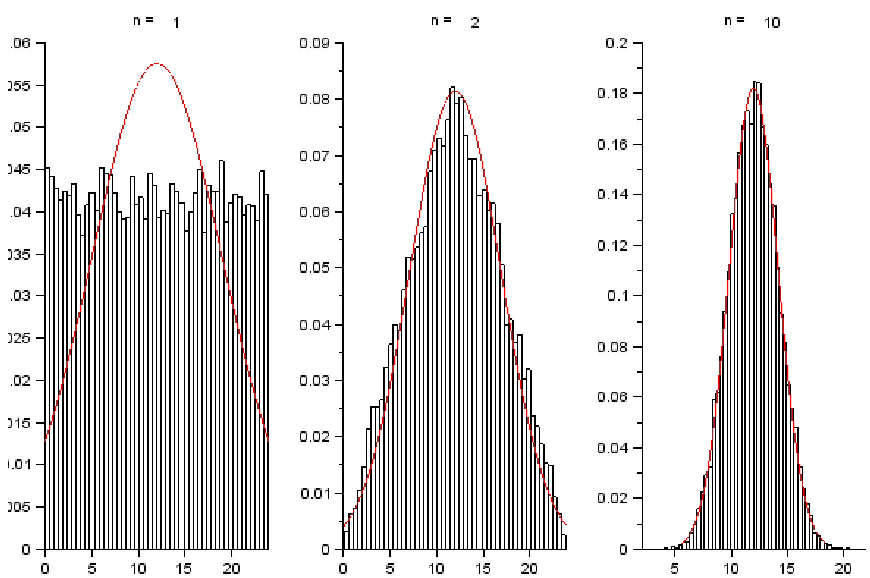
> *Můžeme tedy pro součet cca třiceti a více náhodných veličin (se stejným rozdělením) používat normální rozdělení $N\left(n\mu, n\rho^2\right)$, kde n je počet sčítaných veličin a μ a ρ^2 jsou střední hodnota a rozptyl jedné veličiny.*

Druhý příklad

Do práce jezdím vlakem. Vlak přijíždí každých 24 minut a já přicházím v náhodný okamžik. Doba čekání má tedy spojité rovnoměrné rozdělení od 0 do 24 minut se střední hodnotou 12 minut a rozptylem 48 minut na druhou.

Takto budu jezdit každý den. Jaká bude průměrná doba čekání po n dnech?

Pomohu si opět simulací ve Scilabu. Nejprve vygeneruji v jednom řádku 10 000 hodnot s příslušným rovnoměrným rozdělením. Takových řádků vygeneruji n a po sloupcích zprůměruji. Získám tedy 10 000 hodnot pro průměrné hodnoty. Vykreslíme histogramy:



Vidíme, že i tentokrát nám získané průměrové rozdělení velmi rychle konverguje k příslušnému Gaussovu rozdělení. Velmi dobrou shodu získáváme již pro $n = 10$.

Parametry Gaussova rozdělení získáme snadno: Střední hodnota průměru je stejná jako střední hodnota jednoho čekání, rozptyl je n-krát menší než rozptyl jednoho čekání.

Můžeme tedy pro průměr cca třiceti a více náhodných veličin (se stejným rozdělením) používat normální rozdělení $N\left(\mu, \frac{\rho^2}{n}\right)$, *kde* n *je počet sčítaných veličin a* μ *a* ρ^2 *jsou střední hodnota a rozptyl jedné veličiny.*

Věta o sčítání a odčítání náhodných veličin

Při sčítání náhodných veličin se sčítají jejich střední hodnoty i rozptyly.

Při odčítání náhodných veličin se odčítají jejich střední hodnoty a sčítají jejich rozptyly.

Věta o průměrování náhodných veličin

Při průměrování náhodných veličin se stejným rozdělením je střední hodnota průměru rovna střední hodnotě jedné průměrované veličiny a rozptyl průměru je n-krát menší než rozptyl jedné průměrované veličiny.

Centrální limitní věta

Nechť má náhodná veličina X_i LIBOVOLNÉ rozdělení.
Nechť veličina Y vzniká jako součet nebo průměr n veličin X_i, kde n je velké (tj. třicet a více).
Pak má veličina Y přibližně Gaussovo rozdělení.