Přednáška 2.

- Transformace hustot pravděpodobnosti
- Centrální limitní věta/teorém (CLT)
- Zákon velkých čísel
- Metoda Monte Carlo
- Statistika

Funkce náhodné veličiny

Uvažujme náhodnou veličinu **X**, a z ní vytvořme náhodnou veličinu **Y** jako $Y = h(X)$.

Nechť **X** je spojitá náhodná veličina, $h : R \to R$ ryze monotónní funkce na množině $X(\Omega)$ a h^{-1} je diferencovatelná. Potom náhodná veličina $Y = h(X)$ má hustotu pravděpodobnosti

$$f_Y(y) = f_X\left(h^{-1}(y)\right)\left|\frac{dh^{-1}}{dy}(y)\right|$$

a navíc

$$F_Y(y) = F_X\left(h^{-1}(y)\right)$$

Příklad: $Y = a + bX,$ kde $X \sim N(0,1)$. Jaká je hustota pravděpodobnosti $f_Y(y)$?

Důležité: Pro střední hodnoty platí: $E[Y] = E[h(X)]$.

- Díky tomuto pravidlu můžeme používat operátor střední hodnoty aniž bychom konkrétně uváděli podle které hustoty pravděpodobnosti (distribuční funkce) se počítá.

Zákon velkých čísel

Uvažujme n nezávislých měření (=náhodných veličin) $X_1, X_2, \dots, X_n$. Náhodnou veličinu $\frac{1}{n}\sum_{i=1}^{n} X_i$ nazveme **průměr.**

Pro $n \to \infty$ označme posloupnost $(X_n)_{n=1}^{\infty}$.

ZVČ:

Nechť $X_1, X_2, \dots, X_n$ je posloupnost i.i.d. náhodných veličin, pro které existuje $\mu = E[X_i]$ a $\sigma^2 = E[(X_i - \mu)^2]$ pro všechna i. Pak platí

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu$$

PŘÍKLAD 1: Gaussovo rozdělení

PŘÍKLAD 2: Cauchyho standardní rozdělení má hustotu pravděpodobnosti

$$f_X(x) = \frac{1}{\pi(1+x^2)}$$

a tedy nemá střední hodnotu ani rozptyl!

Centrální limitní věta

Často nás zajímá jaké má rozdělení náhodná veličina, která vznikne **součtem** nebo průměrem náhodných veličin. To ale obecně vůbec není snadné. **CLT** nám říká, za jakých podmínek lze toto rozdělení aproximovat **normálním rozdělením**.

CLT:

Nechť $(X_n)_{n=1}^{\infty}$ *je posloupnost i.i.d. náhodných veličin se společnou střední hodnotou* μ *a konečným rozptylem* σ^2*. Potom platí*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \sim N(0,1)$$

Metoda Monte Carlo

- Algoritmy/metody pro simulace systémů s náhodnými daty podle konkrétního náhodného modelu
- Základní cíl: určení střední hodnoty náhodné veličiny, která je výsledkem náhodného děje.
- Generováním pseudonáhodných čísel se vytvoří potřebný počet pozorování, ze kterého se vypočte vybraná statistika.

Příklady:

- Výpočet určitého integrálu.
- Simulace experimentů s použitím metod Markov-Chain Monte Carlo (MCMC).
- Odhad hustoty pravděpodobnosti (histogram rozestupů).

Úskalí:

- Nutným předpokladem je **dokonalá korespondence mezi modelem a způsobem provedení simulace**. Chyby vedou k nesprávným závěrům („vždycky něco vyjde“)!

Statistika

Definice:

N-tici i.i.d. náhodných veličin $X_1, X_2, \dots, X_n$ *s distribuční funkcí* F *nazýváme* ***náhodný výběr*** *z rozdělení* F.

Definice:

Uvažujme náhodný výběr $X_1, X_2, \dots, X_n$ *.* ***Statistikou*** *nazveme každou funkci náhodného výběru, která nezávisí na konkrétních hodnotách příslušného rozdělení.*

Příklady:

- **Výběrový průměr:** $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- **Výběrový rozptyl:** $s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2$ **,** s_n pak nazveme **výběrová směrodatná odchylka.**

Cvičení 2

1. Nechť $X \sim N(\mu, \sigma^2)$, jakou hustotu má náhodná veličina $Y = \frac{X-\mu}{\sigma}$?
2. Ukažte, že když X a Y jsou nezávislé náhodné veličiny, pak $E[XY] = E[X] \cdot E[Y]$.
3. Aplikujte CLT pro součet náhodných veličin $X_1, X_2, \dots, X_n \sim U(0,1)$ a vykreslete histogram výsledků (využijte metodu Monte Carlo).