Centrální limitní věta

9. cvičení

1. Délky telefonních hovorů paní Zuzany (v min) jsou stejně rozdělené navzájem nezávislé náhodné veličiny se střední hodnotou 5 a rozptylem 25. Tento měsíc paní Zuzana provedla 50 hovorů. S jakou pravděpodobností volala dohromady méně než 280 minut, což je objem volných minut na jejím tarifu?

2. Na server má přístup 100 uživatelů. Z dřívějších zkušeností víme, že uživatel má na serveru průměrně 1200MB dat, směrodatná odchylka množství dat je 400 MB. Jak velký diskový prostor potřebujeme, aby s pravděpodobností 99% nedošlo k jeho zaplnění?

3. Pojišťovna má pojištěno 1 000 osob stejného věku. Pravděpodobnost úmrtí v daném roce je u každého pojištěného 0,01. Pojištěnci platí roční pojistné 1 200 Kč a v případě úmrtí je oprávněné osobě vyplaceno 80 000 Kč.
   (a) Jaký je v daném roce očekávaný zisk pojišťovny?
   (b) Jaká je pravděpodobnost, že pojišťovna utrpí v daném roce ztrátu?

4. Stroj generuje náhodné číslice 0-9, každé číslo se stejnou pravděpodobností.
   (a) Kolik nejméně číslic musíme vygenerovat, abychom s pravděpodobností alespoň 0.975 dostali alespoň jedno sudé číslo? (Spočítejte přesně.)
   (b) Kolik nejméně číslic musíme vygenerovat, abychom s pravděpodobností alespoň 0.975 dostali alespoň 20 sudých. (Použijte aproximaci pomocí CLV.)

5. Házíme pravidelnou symetrickou mincí, přičemž ν_n značí relativní četnost líců v n hodech, tj. $\nu_n = (\#\text{líců})/n$. Kolik musíme provést hodů, aby pravděpodobnost jevu $[|\nu_n - 1/2| \leq 0.1]$ byla alespoň 0.95?
   Řešte pomocí CLV a srovnejte s výsledkem získaným na základě Čebyševovy nerovnosti.

6. Pořádáte narozeninovou oslavu pro 100 hostů a zajímá Vás, kolik máte objednat chlebíčků. Ze zkušenosti víte, že počet chlebíčků, který sní náhodný host, se řídí Poissonovým rozdělením se střední hodnotou 3. Kolik přibližně musíte objednat chlebíčků, chcete-li mít jistotu, že s pravděpodobností 0.95 nebude žádný host hladovět?

Opakování z přednášky

Centrální limitní věta (CLV): Nechť $\{X_n\}$ je posloupnost nezávislých stejně rozdělených náhodných veličin s $0 < \mathsf{Var}\, X_1 < \infty$. Pak

$$\mathsf{P}\left(\frac{\sum_{k=1}^{n} X_k - n\mathsf{E}X_1}{\sqrt{n\,\mathsf{Var}\, X_1}} \leq x\right) \stackrel{n\to\infty}{\longrightarrow} \Phi(x), \quad x \in \mathbb{R},$$

neboli ekvivalentně

$$\mathsf{P}\left(\sqrt{n}\cdot\frac{\overline{X}_n - \mathsf{E}X_1}{\sqrt{\mathsf{Var}\, X_1}} \leq x\right) \stackrel{n\to\infty}{\longrightarrow} \Phi(x), \quad x \in \mathbb{R},$$

kde Φ je distribuční funkci normálního rozdělení $\mathsf{N}(0,1)$, jejíž hodnoty je možné najít např. v tabulkách.
Zkráceně píšeme

$$Z_n := \frac{\sum_{k=1}^{n} X_k - n\mathsf{E}X_1}{\sqrt{n\mathsf{Var}\, X_1}} = \sqrt{n}\cdot\frac{\overline{X}_n - \mathsf{E}X_1}{\sqrt{\mathsf{Var}\, X_1}} \stackrel{asympt.}{\sim} \mathsf{N}(0,1)$$

a říkáme, že Z_n má asymptoticky normální rozdělení.

CLV nám tedy říká, že distribuční funkce F_n veličiny Z_n se při $n \to \infty$ blíží k Φ. Pro n dost velké tedy lze uvažovat

$$\mathsf{P}\left(Z_n \leq x\right) \doteq \Phi(x).$$

Tabulka distribuční funkce a kvantilové funkce $\mathsf{N}(0,1)$

x	0.000	0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450
$\Phi(x)$	0.500	0.520	0.540	0.560	0.579	0.599	0.618	0.637	0.655	0.674
x	0.500	0.550	0.600	0.650	0.700	0.750	0.800	0.850	0.900	0.950
$\Phi(x)$	0.691	0.709	0.726	0.742	0.758	0.773	0.788	0.802	0.816	0.829
x	1.000	1.050	1.100	1.150	1.200	1.250	1.300	1.350	1.400	1.450
$\Phi(x)$	0.841	0.853	0.864	0.875	0.885	0.894	0.903	0.911	0.919	0.926
x	1.500	1.550	1.600	1.650	1.700	1.750	1.800	1.850	1.900	1.950
$\Phi(x)$	0.933	0.939	0.945	0.951	0.955	0.960	0.964	0.968	0.971	0.974
x	2.000	2.050	2.100	2.150	2.200	2.250	2.300	2.350	2.400	2.450
$\Phi(x)$	0.977	0.980	0.982	0.984	0.986	0.988	0.989	0.991	0.992	0.993
x	2.500	2.550	2.600	2.650	2.700	2.750	2.800	2.850	2.900	2.950
$\Phi(x)$	0.994	0.995	0.995	0.996	0.997	0.997	0.997	0.998	0.998	0.998

α	0.5	0.8	0.9	0.95	0.975	0.99	0.995
q_α	0	0.842	1.282	1.645	1.960	2.326	2.576