Příklad (2.1)
Spočítejte rozptyl Beta rozdělení s parametry $a, b > 0$.

Řešení
Využijeme toho, že $\operatorname{var} X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$. Z přednášky víme, že $\mathbb{E}X = \frac{a}{a+b}$. $\mathbb{E}(X^2)$ spočítáme pomocí hustoty $f_X(s) = \frac{1}{B(a,b)} s^{a-1} \cdot (1-s)^{b-1} \cdot 1_{[0,1]}(s)$ jako:

$$\int_{\mathbb{R}} s^2 f_X(s) ds = \frac{1}{B(a,b)} \cdot \int_0^1 s^{a+1}(1-s)^{b-1} ds = \frac{B(a+2,b)}{B(a,b)} = \frac{a+1}{a+1+b}\frac{a}{a+b} \cdot \frac{B(a,b)}{B(a,b)}.$$

$$\operatorname{var} X = \frac{a(a+1)}{(a+b)(a+b+1)} - \left(\frac{a}{a+b}\right)^2 = \frac{a(a+1)(a+b) - a^2(a+b+1)}{(a+b+1)\cdot(a+b)^2} =$$

$$= \frac{a \cdot b}{(a+b+1)\cdot(a+b)^2}.$$

Příklad (2.2)
Negativně binomické rozdělení je diskrétní rozdělení s parametry $r \in \mathbb{N}$ a $p \in (0,1)$. Negativně binomicky rozdělená náhodná veličina X nabývá hodnot z $\mathbb{N}_0$ s pravděpodobnostmi

$$P_{\mathrm{N\,Bi}(r,p)}(k) = P(X = k) = \binom{r-1+k}{r-1} \cdot p^r \cdot (1-p)^k, \qquad k \in \mathbb{N}_0.$$

Odpovídá rozdělení počtu neúspěchů před r-tým úspěchem v sérii nezávislých pokusů s pravděpodobností úspěchu p. Je to vlastně rozdělení doby čekání na úspěch v diskrétním případě.

Označme $F_{\Gamma(\alpha,r)}$ distribuční funkci Gamma rozdělení s parametry $r \in \mathbb{N}$, $\alpha > 0$.

Buď $r \in \mathbb{N}$, $\alpha, t > 0$, $\{p_n\}_{n=1}^{\infty}$ posloupnost čísel z $(0,1)$ splňující $n \cdot p_n \to \alpha$ a $\{t_n\}_{n=1}^{\infty}$ posloupnost čísel ze $\mathbb{Z}_+$ splňující $t_n/n \to t$. Ukažte, že

$$F_{\Gamma(\alpha,r)}(t) = \lim_{n\to\infty} P_{\mathrm{N\,Bi}(r,p_n)}(\{0,\dots,t_n\}).$$

Interpretujte tento výsledek jako výrok o rozdělení dob čekání v diskrétním a spojitém modelu.

Důkaz
Nejprve podle nápovědy ukážeme $P_{\mathrm{N\,Bi}(r,p)}(\{0,\dots,m\}) = P_{\mathrm{Bi}(r+m,p)}(\{r,\dots,r+m\})$. K tomu se podíváme na pravděpodobnostní prostor X reprezentující $r+m$ náhodných pokusů s pravděpodobností úspěchu p. Podle zadání (nebo podle tvaru $P_{\mathrm{N\,Bi}(r,p)}(k)$) víme, že $P_{\mathrm{N\,Bi}(r,p)}(k)$ je pravděpodobnost P množiny těch jevů, kde v prvních $k+r-1$ pokusech nastalo $r-1$ úspěchů a $(r+k)$-tý pokus byl úspěšný (a další dopadly libovolně). Na druhou stranu $P_{\mathrm{Bi}(r+m,p)}(r+k)$ je pravděpodobnost P množiny těch jevů z X, kde je právě $r+k$ úspěchů.

Jelikož pravděpodobnosti jsou konečně aditivní, tak $P_{\mathrm{N\,Bi}(r,p)}(\{0,\dots,m\})$ je součet pravděpodobností $P_{\mathrm{N\,Bi}(r,p)}(k), k \in [m]_0$, tedy pravděpodobnost P množiny těch jevů z X, kde nastalo alespoň r úspěchů, neboť víme, že r-tý pokus někde nastal a poslední pokus, kde nás ještě zajímá r-tý úspěch, je $r+m$-tý. Stejně tak $P_{\mathrm{Bi}(r+m,p)}(\{r,\dots,r+m\})$ je pravděpodobnost P množiny všech jevů z X, kde nastalo alespoň r úspěchů, neboť je to právě $r+k$ pokusů, kde $r+k$ jde od r do $r+\ m$. Tudíž $P_{\mathrm{N\,Bi}(r,p)}(\{0,\dots,m\}) = P_{\mathrm{Bi}(r+m,p)}(\{r,\dots,r+m\})$, neboť jsou obě strany rovny pravděpodobnosti P stejného jevu $\subseteq X$.

Tudíž chceme dokázat $F_{\Gamma(\alpha,r)}(t) = \lim_{n\to\infty} P_{\mathrm{Bi}(r+t_n,p_n)}(\{r,\dots,r+t_n\})$. První věc, které si můžeme všimnout, je

$$P_{\mathrm{Bi}(r+t_n,p_n)}(\{r,\dots,r+t_n\}) = 1 - P_{\mathrm{Bi}(r+t_n,p_n)}(\{0,\dots,r-1\}),$$

neboť pravděpodobnost, že vybereme více než $r+t_n$ nebo méně než 0 prvků z $r+t_n$, je nulová. Z prosemináře navíc víme, že

$$F_{\Gamma(\alpha,r)}(t) = 1 - P_{\mathrm{Poiss}(\alpha t)}(\{0,\dots,r-1\}).$$

Také víme, že

$$P_{\mathrm{Poiss}(\alpha t)}(k) = \lim_{\substack{m\to\infty,\\ p_m\cdot m\to\alpha t}} P_{\mathrm{Bi}(m,p_m)}(k).$$

Z toho můžeme vybrat podposloupnost $m_n = t_n + r$ pro $t_n/n \to t$, tj. $t_n \to \infty$, a označíme $p_{m_n} = p_n$:

$$P_{\mathrm{Poiss}(\alpha t)}(k) = \lim_{\substack{n\to\infty,\\ p_n\cdot(t_n+r)\to\alpha t,\\ t_n/n\to t}} P_{\mathrm{Bi}(t_n+r,p_n)}(k).$$

Jelikož $p_n \to 0$, tak $p_n \cdot r \to 0$, tedy (po troše rozmýšlení, že limita p_n musí existovat) $p_n \cdot t_n \to \alpha \cdot t$ a $p_n \cdot n \to \alpha$, jelikož $t_n/n \to t$ (a limita $p_n \cdot n$ existuje). Teď už to můžeme všechno poskládat a dostaneme

$$F_{\Gamma(\alpha,r)}(t) = 1 - P_{\mathrm{Poiss}(\alpha t)}(\{0,\dots,r-1\}) = 1 - \lim_{\substack{n\to\infty,\\ p_n\cdot(t_n+r)\to\alpha t,\\ t_n/n\to t}} P_{\mathrm{Bi}(t_n+r,p_n)}(\{0,\dots,r-1\}) =$$

$$= \lim_{\substack{n\to\infty,\\ p_n\cdot(t_n+r)\to\alpha t,\\ t_n/n\to t}} 1 - P_{\mathrm{Bi}(t_n+r,p_n)}(\{0,\dots,r-1\}) = \lim_{\substack{n\to\infty,\\ p_n\cdot(t_n+r)\to\alpha t,\\ t_n/n\to t}} P_{\mathrm{Bi}(r+t_n,p_n)}(\{r,\dots,r+t_n\}).$$

□

Řešení (Interpretace)
Na levé straně je pravděpodobnost, že do času t nastala r-tá z událostí, které nastávají v libovolném (spojitém) čase tak, že jich v jednotkovém čase nastane ve střední hodnotě α (počet má Possonovo rozdělení, tedy limitní binomické – jako bychom dělali v každém časovém okamžiku pokus).

Pravou stranu si můžeme představit tak, že (limitně) každý „jednotkový“ okamžik do času t namodelujeme n pokusy (jelikož $t_n/n \to t$, takže vlastně $t_n \approx t \cdot n$). V každém takovém pokusu máme pravděpodobnost (limitně) α/n. Z četnostního pojetí pravděpodobnosti to tedy znamená, že za každý jednotkový čas se nám stane α událostí a my se snažíme namodelovat to, aby se mohli stát „kdykoliv“ během toho úseku. A pravá strana pak vyjadřuje pravděpodobnost, že se r-tá událost stala do času t.

Tedy je to vlastně totéž, v obou případech se díváme zda r-tá událost nastala do času t s tím, že každá událost má šanci se stát „kdykoliv“.

Příklad (2.3)
Buď $T_{r:n}$ čas r-té události z n nezávisle rovnoměrně rozdělených časů v $(0,1)$, tj. $T_{r:n}$ má $\text{Beta}(r, n-r+1)$ rozdělení.

Buď $\{s_n\}_{n=1}^{\infty}$ posloupnost čísel z $(0,\infty)$ splňující $n/s_n \to \alpha > 0$. Ukažte, že pro každé $r \in \mathbb{N}$ a $t > 0$ platí

$$F_{\Gamma(\alpha,r)}(t) = \lim_{n\to\infty} P(s_n T_{r:n} \leq t).$$

Interpretujte tento výsledek jako výrok o rozdělení časů událostí na $(0,\infty)$.

Řešení
Přepíšeme rovnost z definic:

$$F_{\Gamma(\alpha,r)}(t) = \int_0^t \frac{\alpha^r x^{r-1} e^{-\alpha x}}{(r-1)!} dx =$$

$$= \lim_{\substack{n\to\infty,\\ n/s_n\to\alpha}} \int_0^{t/s_n} \frac{1}{B(r,n-r+1)} s^{r-1}(1-s)^{n-r} ds = \lim_{\substack{n\to\infty,\\ n/s_n\to\alpha}} P(s_n T_{r:n} \le t).$$

Podle věty o substituci pro $z = s \cdot s_n$ dostaneme

$$\int_0^t \frac{\alpha^r x^{r-1} e^{-\alpha x}}{(r-1)!} dx = \lim_{\substack{n\to\infty,\\ n/s_n\to\alpha}} \int_0^t \frac{1}{B(r,n-r+1)} \left(\frac{z}{s_n}\right)^{r-1} \left(1-\frac{z}{s_n}\right)^{n-r} \frac{dz}{s_n}.$$

Protože $n/s_n \to \alpha$ a $\frac{1}{B(r,n-r+1)} < n^{r-1}$ tak se dají funkce uvnitř pravého integrálu omezit (až na nějaké první členy) konstantou, tedy podle Lebesgueovy věty můžeme prohodit integrál a limitu. Následně pokud ukážeme rovnost vnitřků integrálů, tak se integrály také rovnají. Tedy dokazujeme

$$\frac{\alpha^r x^{r-1} e^{-\alpha x}}{(r-1)!} = \lim_{\substack{n\to\infty,\\ n/s_n\to\alpha}} \frac{1}{B(r,n-r+1)} \left(\frac{x}{s_n}\right)^{r-1} \left(1-\frac{x}{s_n}\right)^{n-r} \frac{1}{s_n}.$$

Rozepíšeme B z definice, vydělíme obě strany rovnosti r a vynásobíme x:

$$\frac{(\alpha \cdot x)^r e^{-\alpha x}}{r!} = \lim_{\substack{n\to\infty,\\ n/s_n\to\alpha}} \frac{n!}{r! \cdot (n-r)!} \left(\frac{x}{s_n}\right)^{r} \left(1-\frac{x}{s_n}\right)^{n-r}.$$

A to je přesně rovnost, kterou jsme dokazovali na prosemináři u Poissonova rozdělení, pro $k = r$ a $\lambda = \alpha \cdot x$.

Řešení (Interpretace)
Levá strana je totéž co v předchozí úloze, jen se hodí vypíchnout ne to, že události probíhají spojitě, ale to, že probíhají „do nekonečna“, zatímco v Beta rozdělení probíhají jen v časech $(0,1)$. Na pravé straně si ale můžeme představit, že místo rozdělení v časech $(0,1)$ máme rozdělení v časech $(0,s_n)$, kde s_n jde do nekonečna, tedy prakticky máme události probíhající „do nekonečna“.

Na levé straně tedy zase za jednotkový okamžik proběhne ve střední hodnotě α událostí (simulovaných nekonečně pokusy). Stejně tak vpravo v časech $(0,s_n)$ máme $s_n \cdot \alpha$ (jelikož $n/s_n \to \alpha$, tedy $n \approx s_n \cdot \alpha$) událostí, tedy na jeden časový okamžik α událostí.

A zase se díváme na pravděpodobnost výskytu r-té události před časem t, v obou případech (vlevo doopravdy, vpravo limitně) na časech $(0,\infty)$ a v obou případech v každém jednotkovém okamžiku proběhne ve střední hodnotě α událostí.