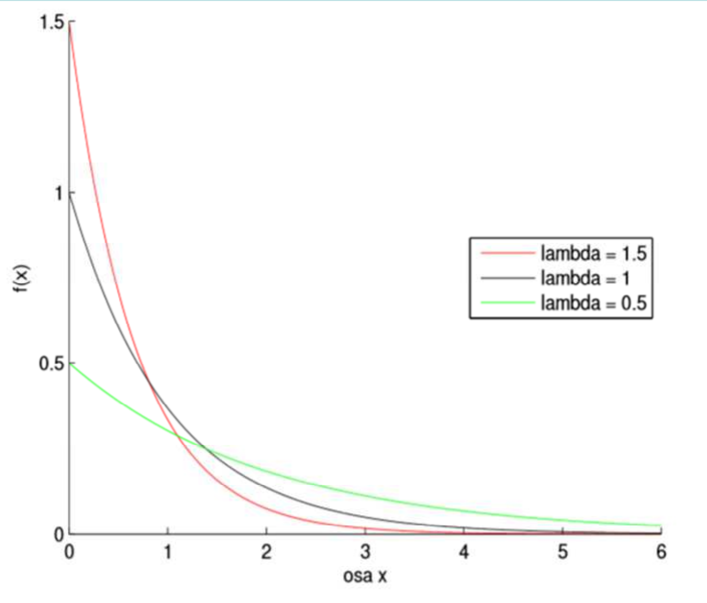
3. Exponenciální rozložení a jeho vlastnosti

3.1. Definice: Spojitá náhodná veličina X má exponenciální rozložení s parametrem $\lambda > 0$, jestliže hustota $\varphi(x)$ má tvar:

$$\varphi(x) = \begin{cases} \lambda e^{-\lambda x} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases}$$

. Zkráceně píšeme $X \sim Ex(\lambda)$.

Průběh hustoty exponenciálního rozložení pro různé hodnoty parametru λ:



3.2. Poznámka: Náhodná veličina X udává dobu čekání na příchod nějaké události, která se může dostavit každým okamžikem se stejnou šancí bez ohledu na dosud pročekanou dobu. (Jde o tzv. čekání bez paměti.) Přitom $1/\lambda$ vyjadřuje střední hodnotu doby čekání. Lze odvodit, že:

a) distribuční funkce $\Phi(x) = \begin{cases} 1 - e^{-\lambda x} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases}$

b) funkce přežití $\Psi(x) = 1 - \Phi(x) = \begin{cases} e^{-\lambda x} & \text{pro } x > 0 \\ 1 & \text{jinak} \end{cases}$

c) intenzita poruchy (též riziková funkce) $\lambda(x) = -\dfrac{\Psi'(x)}{\Psi(x)} = -\dfrac{-\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$

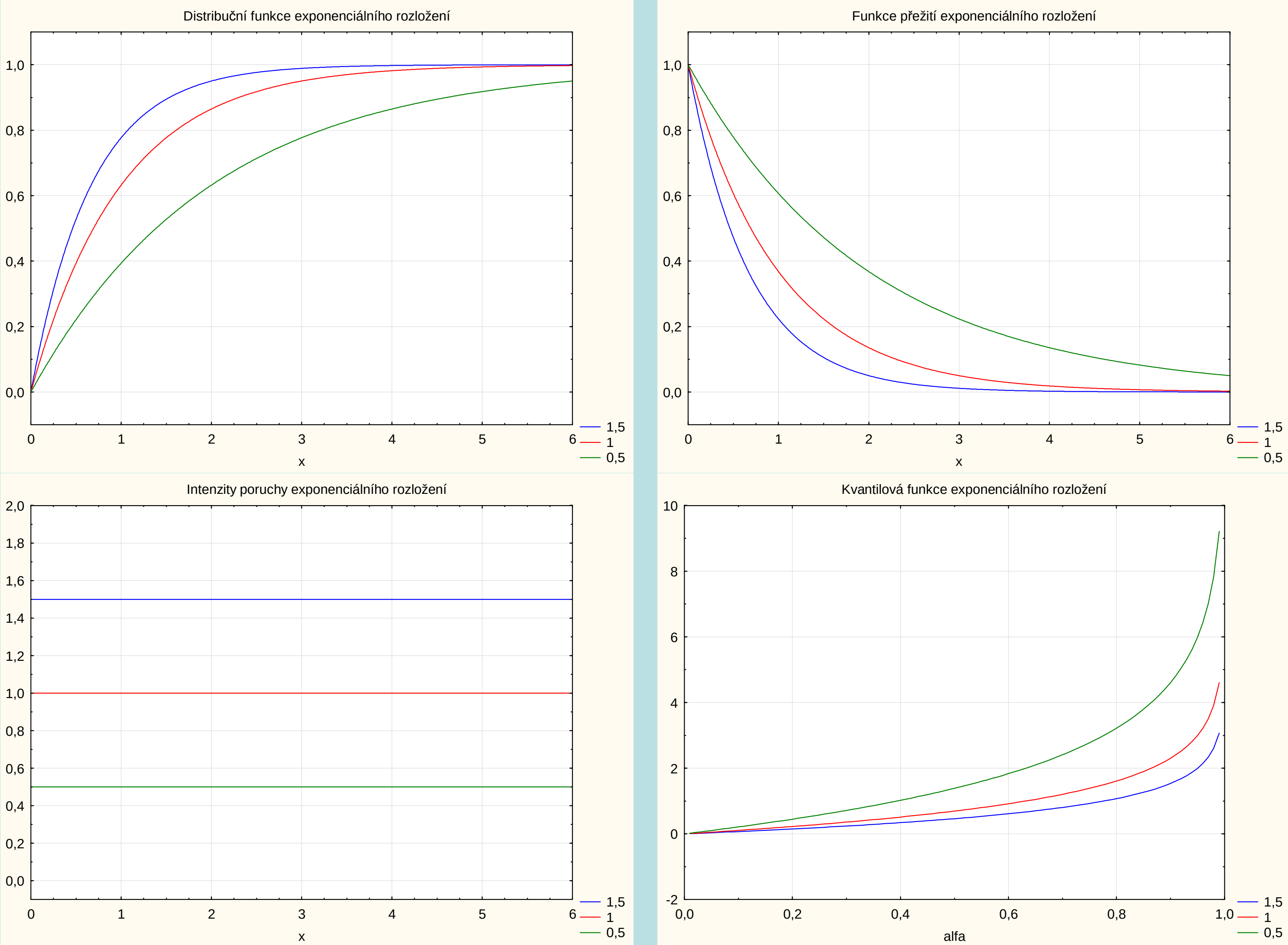
d) kvantilová funkce $\Phi^{-1}(\alpha) = -\dfrac{1}{\lambda}\ln(1-\alpha),\ \alpha \in (0,1)$

e) střední hodnota $E(X) = \dfrac{1}{\lambda}$

f) rozptyl $D(X) = \dfrac{1}{\lambda^2}$

g) medián $x_{0,50} = \dfrac{\ln 2}{\lambda}$

Průběhy důležitých funkcí exponenciálního rozložení



3.3. Poznámka: Exponenciální rozložení Ex(λ) je speciálním případem dvouparametrického exponenciálního rozložením Ex(A, λ), kde parametr A > 0 udává dobu, po kterou událost nemůže nastat. Hustota:

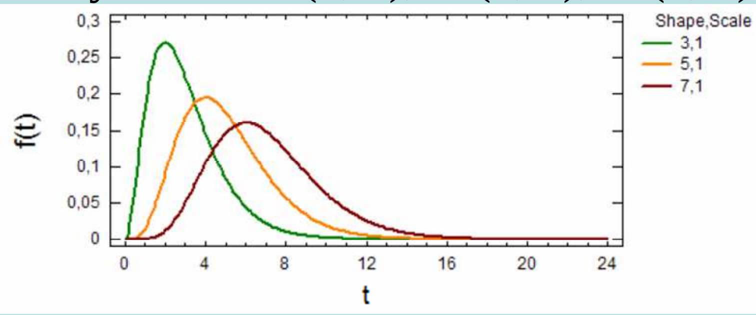
$$\varphi(x) = \begin{cases} \lambda e^{-\lambda(x-A)} \text{ pro } x > A \\ 0 \text{ jinak} \end{cases}.$$

Exponenciální rozložení je speciálním případem Erlangova rozložení Er(k, λ) pro k = 1. Náhodná veličina X s rozložením Er(k, λ) vyjadřuje souhrnnou dobu čekání na k-tý výskyt události, která se může dostavit každým okamžikem se stejnou šancí bez ohledu na dosud pročekanou dobu. Přitom $\frac{1}{\lambda}$ je střední hodnota doby čekání od výskytu předešlé události.

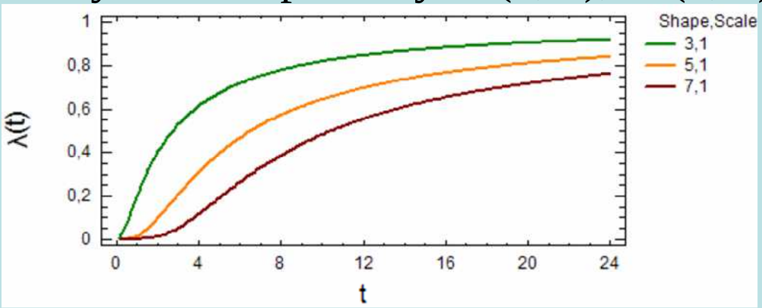
Hustota: $$\varphi(x) = \begin{cases} \frac{(\lambda x)^{k-1}}{(k-1)!} \lambda e^{-\lambda x} \text{ pro } x > 0 \\ 0 \text{ jinak} \end{cases}.$$

(Erlangovo rozložení je speciálním případem gama rozložení, kde první parametr je přirozené číslo.)

Grafy hustot Er(3, 1), Er(5, 1), Er(7, 1):



Grafy intenzit poruchy Er(3, 1), Er(5, 1), Er(7, 1):



Intenzita poruchy Erlangova rozložení je rostoucí funkce, proto je toto rozložení vhodné pro modelování procesů stárnutí.

3.4. Příklad (Praktické využití základních vlastností exponenciálního rozložení)

Dlouhodobým pozorováním v určité prodejně bylo zjištěno, že 40 % zákazníků je obslouženo do 3 minut. Lze předpokládat, že doba čekání se řídí exponenciálním rozložením.

a) Určete parametr λ exponenciálního rozložení.

b) Vypočtěte střední hodnotu doby čekání na obsluhu.

c) Jaká je doba čekání, kterou polovina osob nepřekročí?

d) Jaké procento zákazníků bude na obsluhu čekat déle než 6 minut?

Řešení:

Ad a) Je známo, že $\Phi^{-1}(0{,}4) = 3$. Přitom $\Phi^{-1}(\alpha) = -\frac{1}{\lambda}\ln(1-\alpha)$, tedy

$$\lambda = -\frac{\ln(1-\alpha)}{\Phi^{-1}(\alpha)} = -\frac{\ln(1-0{,}4)}{3} = 0{,}1703$$

Ad b) $E(X) = \frac{1}{\lambda} = \frac{1}{0{,}1703} = 5{,}87\,\text{min} = 5\,\text{min}\,52\text{s}$

Ad c) Hledáme medián, tedy počítáme $\frac{\ln 2}{\lambda} = \frac{\ln 2}{0{,}1703} = 4{,}07\,\text{min} = 4\,\text{min}\,4\text{s}$

Ad d) $P(X > 6) = \Psi(6) = e^{-\lambda \cdot 6} = e^{-0{,}1703 \cdot 6} = 0{,}36$.

Znamená to, že 36 % zákazníků bude čekat déle než 6 minut.

3.5. Věta: Nechť $X \sim Ex(\lambda)$. Pak platí: $\forall t > 0, \forall h > 0 : P(X > t + h / X > t) = P(X > h)$

Vysvětlení: Tato věta vysvětluje, proč se exponenciálnímu rozložení říká rozložení bez paměti. Jestliže náhodná veličina X udává dobu do poruchy nějakého zařízení, pak pravděpodobnost, že zařízení, které pracovalo po dobu aspoň t, bude pracovat bez poruchy aspoň po dobu t + h, je stejná jako pravděpodobnost, že zařízení bude pracovat bez poruchy po dobu aspoň h – jako kdyby „zapomnělo“ již odpracovanou dobu t.

Důkaz:

$$P(X > t + h / X > t) = \frac{P(X > t + h \wedge X > t)}{P(X > t)} = \frac{P(X > t + h)}{P(X > t)} = \frac{\Psi(t + h)}{\Psi(t)} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda h} =$$

$$= \Psi(t) = P(X > h)$$

3.6. Příklad: Výrobce žárovek udává, že průměrná doba životnosti jeho žárovek je 10 000 h. V rámci své propagační kampaně chce garantovat dobu t, do níž s spálí nejvýše 3 % žárovek. Stanovte tuto dobu za předpokladu, že životnost žárovky se řídí exponenciálním rozložením.

Řešení: $X \sim Ex(\lambda)$, $E(X) = \frac{1}{\lambda} = 10000 \Rightarrow \lambda = \frac{1}{10000}$. Hledáme t tak, aby platilo:

$$0{,}03 = P(X \le t) = \Phi(t) = 1 - e^{-\frac{t}{10000}} \Rightarrow t = -10000 \cdot \ln 0{,}97 = 304{,}6\ \text{h}$$

3.7. Věta: Nechť $X \sim Ex(\lambda)$. Pak transformovaná náhodná veličina Y = λX ~ $Ex(1)$.
(Rozložení $Ex(1)$ se nazývá standardizované exponenciální rozložení.)

Důkaz: $\Phi_*(y) = P(Y \le y) = P(\lambda X \le y) = P\left(X \le \frac{y}{\lambda}\right) = \Phi\left(\frac{y}{\lambda}\right) = \begin{cases} 1 - e^{-y} & \text{pro } y > 0 \\ 0 & \text{jinak} \end{cases}$, tedy Y ~ $Ex(1)$.

3.8. Věta: Nechť X ~ Rs(0, 1). Pak transformovaná náhodná veličina Y = -ln X ~ $Ex(1)$.
Důkaz:

$$\Phi_*(y) = P(Y \le y) = P(-\ln X \le y) = P\left(X \ge -e^{-y}\right) = 1 - P\left(X \le -e^{-y}\right) = 1 - \Phi\left(e^{-y}\right) = \begin{cases} 1 - e^{-y} & \text{pro } y > 0 \\ 0 & \text{jinak} \end{cases}$$

3.9. Poznámka: Vět 3.7. a 3.8. se využívá při generování realizací náhodné veličiny Y ~ $\mathrm{Ex}(\lambda)$ na počítači: vygenerujeme n náhodných čísel $x_1, \ldots, x_n$ ~ Rs(0, 1) a transformujeme je vztahem $y_i = -\frac{1}{\lambda}\ln(1 - x_i),\ i = 1, \ldots, n$. Čísla $y_1, \ldots, y_n$ jsou realizace náhodné veličiny Y ~ $\mathrm{Ex}(\lambda)$.

3.10. Věta: Nechť X_1, X_2 jsou stochasticky nezávislé náhodné veličiny, $X_i \sim \mathrm{Ex}(\lambda_i)$, i = 1, 2. Pak transformovaná náhodná veličina $Y = \min\{X_1, X_2\} \sim \mathrm{Ex}(\lambda_1 + \lambda_2)$.

Důkaz:

$$\Phi_*(y) = P(Y \le y) = P(\min\{X_1, X_2\} \le y) = P(X_1 \le y \vee X_2 \le y) = 1 - P(X_1 > y)P(X_2 > y) =$$

$$= 1 - \Psi_1(y)\Psi_2(y) = \begin{cases} 1 - e^{-\lambda_1 y}e^{-\lambda_2 y} = 1 - e^{-(\lambda_1 + \lambda_2)y} \text{pro } y > 0 \\ 0 \text{ jinak} \end{cases},$$

tedy $Y \sim \mathrm{Ex}(\lambda_1 + \lambda_2)$.

3.11. Poznámka: Tvrzení věty 3.10. lze zobecnit i na n stochasticky nezávislých veličin $X_1, \ldots, X_n$, $X_i \sim \mathrm{Ex}(\lambda_i)$, i = 1, …, n. Pak transformovaná náhodná veličina $Y = \min\{X_1, \ldots, X_n\} \sim \mathrm{Ex}(\lambda_1 + \ldots + \lambda_n)$.

3.12. Věta: Nechť X_1, X_2 jsou stochasticky nezávislé náhodné veličiny, $X_i \sim Ex(\lambda)$, i = 1, 2. Pak transformovaná náhodná veličina $Y = X_1 + X_2 \sim Er(2, \lambda)$.

Důkaz:

Podle věty o konvoluci dostáváme:

$$\varphi_*(y) = \int_{-\infty}^{\infty} \varphi_1(x_1)\varphi_2(y - x_1)dx_1 = \left|x_1 > 0, y - x_1 > 0 \Rightarrow 0 < x_1 < y\right| = \int_0^y \lambda e^{-\lambda x_1} \lambda e^{-\lambda(y - x_1)} dx_1 =$$

$$= \lambda^2 e^{-\lambda y} \int_0^y dx_1 = \lambda^2 y e^{-\lambda y} \text{ pro } y > 0, = 0 \text{ jinak}$$

Přitom rozložení Er(2, λ) má hustotu $\varphi(x) = \begin{cases} \dfrac{(\lambda x)^{2-1}}{(2-1)!} \lambda e^{-\lambda x} = \lambda^2 x e^{-\lambda x} & \text{pro } x > 0 \\ 0 \text{ jinak} & \end{cases}$.

3.13. Poznámka: Tvrzení věty 3.12. lze zobecnit i na n stochasticky nezávislých veličin $X_1, \ldots, X_n$, $X_i \sim Ex(\lambda)$, i = 1, …, n. Pak transformovaná náhodná veličina $Y = \sum_{i=1}^{n} X_i \sim Er(n, \lambda)$.

3.14. Příklad: Zákazník prochází třemi nezávislými stanicemi obsluhy, přičemž v každé z nich se doba obsluhy řídí exponenciálním rozložením se střední hodnotou 1 minuta. Jaká je pravděpodobnost, že celková doba obsluhy nepřesáhne 2 minuty?

Řešení: Označme X_i dobu obsluhy na i-té stanici, $X_i \sim Ex(1)$, i = 1, 2, 3 a $Y = X_1 + X_2 + X_3$ celkovou dobu obsluhy. Podle poznámky 3.13. se celková doba obsluhy řídí rozložením Er(3,1), tedy hustota $\varphi(y) = \frac{y^2}{2} e^{-y}$ pro $y > 0$. Počítáme:

$$P(Y \le 2) = \int_0^2 \frac{y^2}{2} e^{-y} dy = \ldots = 1 - 5e^{-2} = 0,3233$$

Pravděpodobnost, že celková doba obsluhy nepřesáhne 2 minuty, je 0,3233.

3.15. Věta: Nechť X_1, X_2 jsou stochasticky nezávislé náhodné veličiny, $X_i \sim Ex(\lambda)$, i = 1, 2. Pak

$$P(X_2 > X_1) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Důkaz: $P(X_2 > X_1) = P((X_1, X_2) \in S)$, kde $S = \{(x_1, x_2) \in R^2; x_1 > 0, x_2 > 0, x_2 > x_1\}$, tedy

$$P((X_1, X_2) \in S) = \iint_S \varphi(x_1, x_2)dx_1dx_2 = \iint_S \varphi_1(x_1)\varphi_2(x_2)dx_1dx_2 = \int_0^\infty \left[\int_{x_1}^\infty \lambda_1 e^{-\lambda_1 x_1} \lambda_2 e^{-\lambda_2 x_2} dx_2\right] dx_1 =$$

$$= \lambda_1 \lambda_2 \int_0^\infty e^{-\lambda_1 x_1} \left[-\frac{1}{\lambda_2} e^{-\lambda_2 x_2}\right]_{x_1}^\infty dx_1 = \lambda_1 \int_0^\infty e^{-\lambda_1 x_1} e^{-\lambda_2 x_1} dx_1 = -\frac{\lambda_1}{\lambda_1 + \lambda_2}\left[e^{-x_1(\lambda_1 + \lambda_2)}\right]_0^\infty = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

3.16. Věta: Nechť X ~ Ex(λ). Pak transformovaná náhodná veličina $Y = 2\lambda X \sim \chi^2(2)$.

Důkaz: Hustota rozložení $\chi^2(n)$ je $\varphi(x) = \begin{cases} \dfrac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \text{ pro } x > 0 \\ 0 \text{ jinak} \end{cases}$. V našem případě n = 2,

tedy $\varphi(x) = \begin{cases} \dfrac{1}{2} e^{-\frac{x}{2}} \text{ pro } x > 0 \\ 0 \text{ jinak} \end{cases}$. Počítáme distribuční funkci veličiny Y:

$$\Phi_*(y) = P(Y \le y) = P(2\lambda X \le y) = P\left(X \le \frac{y}{2\lambda}\right) = \Phi\left(\frac{y}{2\lambda}\right) = 1 - e^{-\frac{y}{2}} \text{ pro } y > 0, = 0 \text{ jinak}$$

Derivováním obdržíme hustotu:

$\varphi_*(y) = \dfrac{d\Phi_*(y)}{dy} = \dfrac{1}{2} e^{-\frac{y}{2}} \text{ pro } y > 0, = 0 \text{ jinak}$, tedy $Y \sim \chi^2(2)$.

3.17. Poznámka: Tvrzení věty 3.16. lze zobecnit i na n stochasticky nezávislých veličin

$X_1, \ldots, X_n$, $X_i \sim Ex(\lambda)$, i = 1, …, n. Pak transformovaná náhodná veličina $Y = 2\lambda \sum_{i=1}^{n} X_i \sim \chi^2(2n)$.

3.18. Věta: Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozložení Ex(λ). Označme $M = \frac{1}{n}\sum_{i=1}^{n} X_i$ výběrový průměr. Pak meze 100(1-α)% intervalu spolehlivosti pro střední hodnotu $\frac{1}{\lambda}$ jsou:

$$D = \frac{2nM}{\chi^2_{1-\alpha/2}(2n)}, \quad H = \frac{2nM}{\chi^2_{\alpha/2}(2n)}.$$

Důkaz: Podle poznámky 3.17. náhodná veličina $Y = 2\lambda nM \sim \chi^2(2n)$. Z definice 100(1-$\alpha$)% intervalu spolehlivosti dostáváme:

$$\forall \lambda > 0 : 1-\alpha = P\left(\chi^2_{\alpha/2}(2n) < 2\lambda nM < \chi^2_{1-\alpha/2}(2n)\right) = P\left(\frac{1}{\chi^2_{1-\alpha/2}(2n)} < \frac{1}{2\lambda nM} < \frac{1}{\chi^2_{\alpha/2}(2n)}\right) =$$

$$= P\left(\frac{2nM}{\chi^2_{1-\alpha/2}(2n)} < \frac{1}{\lambda} < \frac{2nM}{\chi^2_{\alpha/2}(2n)}\right)$$

3.19. Příklad: V jisté prodejně potravin bylo na základě náhodného výběru 50 zákazníků zjištěno, že průměrná doba obsluhy u pokladny je 30 s. Předpokládejme, že doba obsluhy je náhodná veličina s exponenciálním rozložením. Najděte 95% empirický interval spolehlivosti pro střední hodnotu doby obsluhy.

Řešení: n = 50, m = 30, α = 0,05

$$d = \frac{2nm}{\chi^2_{1-\alpha/2}(2n)} = \frac{2 \cdot 50 \cdot 30}{\chi^2_{0,975}(100)} = \frac{3000}{129,501} = 23,16$$

$$h = \frac{2nm}{\chi^2_{\alpha/2}(2n)} = \frac{2 \cdot 50 \cdot 30}{\chi^2_{0,025}(100)} = \frac{3000}{74,222} = 40,42$$

Střední hodnota doby obsluhy se s pravděpodobností aspoň 0,95 nachází v intervalu 23 s až 40 s.

3.20. Poznámka: Pro větší rozsahy náhodných výběrů ($n \geq 30$) lze pro střední hodnotu $\frac{1}{\lambda}$ použít asymptotický 100(1-α)% interval spolehlivosti založený na centrální limitní větě.

Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozložení Ex(λ), $M = \frac{1}{n}\sum_{i=1}^{n} X_i$ je výběrový průměr. Pak střední hodnota $E(M) = \frac{1}{\lambda}$ a rozptyl $D(M) = \frac{1}{n\lambda^2}$. Standardizací výběrového průměru dostaneme veličinu $U = \frac{M - E(M)}{\sqrt{D(M)}} = \frac{M - \frac{1}{\lambda}}{\sqrt{\frac{1}{n\lambda^2}}} = \frac{\left(M - \frac{1}{\lambda}\right)\sqrt{n}}{\frac{1}{\lambda}} \approx N(0,1)$. Konvergence k rozložení N(0,1) se neporuší, když $\frac{1}{\lambda}$ ve jmenovateli nahradíme M, tedy $U = \frac{\left(M - \frac{1}{\lambda}\right)\sqrt{n}}{M} \approx N(0,1)$.

$$\forall \lambda > 0 : 1 - \alpha \leq P\left(-u_{1-\alpha/2} < \frac{\left(M - \frac{1}{\lambda}\right)\sqrt{n}}{M} < u_{1-\alpha/2}\right) = P\left(M - \frac{M}{\sqrt{n}}u_{1-\alpha/2} < \frac{1}{\lambda} < M + \frac{M}{\sqrt{n}}u_{1-\alpha/2}\right)$$, tedy

$$D = M - \frac{M}{\sqrt{n}}u_{1-\alpha/2}, \; H = M + \frac{M}{\sqrt{n}}u_{1-\alpha/2}.$$

3.21. Příklad: Pro údaje z příkladu 3.19. spočtěte meze 95% intervalu spolehlivosti pro střední hodnotu doby obsluhy podle vzorů uvedených v poznámce 3.20.

Řešení: n = 50, m = 30, α = 0,05

$$d = m - \frac{m}{\sqrt{n}} u_{1-\alpha/2} = 30 - \frac{30}{\sqrt{50}} 1{,}96 = 21{,}68$$

$$h = m + \frac{m}{\sqrt{n}} u_{1-\alpha/2} = 30 + \frac{30}{\sqrt{50}} 1{,}96 = 38{,}32$$

Střední hodnota doby obsluhy se s pravděpodobností aspoň 0,95 nachází v intervalu 22 s až 38 s.

3.22. Poznámka: Výhody a nevýhody exponenciálního rozložení v praxi
Výhoda: Exponenciálního rozložení má jednoduché vyjádření hustoty a jeho hustota má průběh, který dobře popisuje řadu reálných dějů.
Nevýhoda: Exponenciální rozložení závisí pouze na jediném parametru, je tedy málo flexibilní. V některých situacích, např. při modelování výše pojistného plnění, špatně modeluje interval nejnižších hodnot a také interval nejvyšších hodnot, protože hustota příliš rychle klesá k nule.

Vzorce pro meze 100(1-α)% intervalu spolehlivosti pro střední hodnotu rozložení Ex(λ)

1. způsob: Využití Pearsonova rozložení chí-kvadrát:

$$d = \frac{2nm}{\chi^2_{1-\alpha/2}(2n)}, \quad h = \frac{2nm}{\chi^2_{\alpha/2}(2n)}$$

2. způsob: Využití standardizovaného normálního rozložení:

$$d = m - \frac{m}{\sqrt{n}} u_{1-\alpha/2}, \quad h = m + \frac{m}{\sqrt{n}} u_{1-\alpha/2}$$