9.1 Moment generující funkce

▷ **Příklad 9.1.** Najděte generující funkci pro náhodnou veličinu X s binomickým rozdělením s parametry n, p a ověřte, že $E(X) = np$, $\sigma_X^2 = np(1-p)$.
Víme, že pro náhodnou veličinu X definovanou v pravděpodobnostním prostoru $(\Omega, \mathcal{F}, P)$ je generující funkce definována jako $M_X(t) = E(e^{tX})$ a že platí $M_X^{(n)}(0) = E(X^n)$. Tedy speciálně pro střední hodnotu náhodné veličiny X dostáváme $E(X) = M_X'(0)$ a pro rozptyl $\sigma_X^2 = E(X^2) - (E(X))^2 = M_X''(0) - (M_X'(0))^2$.
Pro náhodnou veličinu s binomickým rozdělením proto platí:

$$\begin{aligned}
M_X(t) &= \sum_{x=0}^{n} e^{tx} f_X(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x} \\
&= (pe^t + q)^n \\
M_X'(t) &= n(pe^t + q)^{n-1} \cdot pe^t \\
M_X'(0) &= n(pe^0 + q)^{n-1} \cdot pe^0 = n(1)p = np \\
M_X''(t) &= n(n-1)(pe^t + q)^{n-2} \cdot pe^t \cdot pe^t + n(pe^t + q)^{n-1} \cdot pe^t \\
&= npe^t[(n-1)(pe^t + q)^{n-2} \cdot pe^t + (pe^t + q)^{n-1}] \\
M_X''(0) &= np[(n-1)(p+q)^{n-2} \cdot p + (p+q)^{n-1}] = np[(n-1)p + 1] \\
\sigma_X^2 &= np[(n-1)p + 1] - (np)^2 = np[np - p + 1 - np] = np(1-p)
\end{aligned}$$

► **Příklad 9.1.** Pro náhodnou veličinu X s geometrickým rozdělením s parametrem p, ($f_X(x) = pq^{x-1}$, kde $x = 1, 2, 3, \ldots$) určete moment generující funkci a pomocí ní určete střední hodnotu a rozptyl. $[M_X(t) = \frac{pe^t}{1-qe^t}, t < \ln q]$

► **Příklad 9.2.** Pro náhodnou veličinu X s Poissonovým rozdělením s parametrem λ ($f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, kde $x = 0, 1, 2, \ldots$) určete moment generující funkci a pomocí ní určete střední hodnotu a rozptyl. $[M_X(t) = exp\{\lambda(e^t - 1)\}]$

9.2 Náhodné veličiny s normálním rozdělením

V následujícím budeme značit Φ distribuční funkci normovaného normálního rozdělení. Hodnotu distribuční funkce v bodě x můžete najít ve statických tabulkách nebo v R za pomoci `pnorm(x)`.

▷ **Příklad 9.2.** Předpokládejte, že dojíždíte každý den z okraje města do centra autem. Jízda Vám v průměru trvá 24 minut se $\sigma = 3.8$ minut. Čas, který Vám jízda zabere je ovlivněn hustotou dopravy, počasím apod. Můžete ale předpokládat, že má normální rozdělení.
Jaká je pravděpodobnost, že Vám cesta do práce bude trvat déle než půl hodiny?
Víme, že $X \sim \mathcal{N}(24, 3.8^2)$. Zajímá nás

$$p = P(\{X > 30\}) = 1 - P(\{X \le 30\}) = 1 - P\Big(\Big\{\frac{X-24}{3.8} \le \frac{30-24}{3.8}\Big\}\Big) = 1 - \Phi\Big(\frac{6}{3.8}\Big)$$

```
> p<-1-pnorm(6/3.8)
> p
[1] 0.05717406
#nebo
> 1-pnorm(30,mean=24,sd=3.8)
[1] 0.05717406
```

Určete pravděpodobnost, že při dalších 3 jízdách do práce, právě 2 budou trvat déle než půl hodiny. Počet jízd, které budou trvat déležnež půl hodiny, během třech jízd, je náhodná veličina Y s binomickým rozdělením s parametry $n = 3$, $p = 0.057$. Zajímá nás $P(\{Y = 2\}) = \binom{3}{2}p^2(1-p) = 0.0092$.

```
> dbinom(2,size=3,prob=p)
[1] 0.009245937
```

▷ **Příklad 9.3.** Zatížení letadla s 200 místy nemá překročit 18 250 kg. Jaká je pravděpodobnost, že při plném obsazení letadla bude tato hodnota překročena, má-li hmotnost cestujícího střední hodnotu 90 a rozptyl 100.
Hmotnost každého cestujícího je náhodná veličina X_i, $i = 1, 2, \cdots 200$, o jejímž rozdělení víme jen to, $\mu_i = 90$ a $\sigma = 10$. Označíme-li si $X = \sum_{i=1}^{n} X_i$, pak nás zajímá $P(\{X > 18\,250\})$. Dle centrální limitní věty platí, že náhodná veličina $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$ má přibližně standardní normální rozdělení $\mathcal{N}(0, 1)$. Proto:

$$\begin{aligned} P(\{X > 18\,250\}) &= 1 - P(\{X \le 18\,250\}) = 1 - P\Big(\Big\{\frac{X - n\mu}{\sqrt{n}\sigma} \le \frac{18\,250 - n\mu}{\sqrt{n}\sigma}\Big\}\Big) \\ &= 1 - P\Big(\Big\{\frac{X - 200 \cdot 90}{\sqrt{200} \cdot 10} \le \frac{18\,250 - 200 \cdot 90}{\sqrt{200} \cdot 10}\Big\}\Big) = 1 - \Phi(1.767767) = 0.03855 \end{aligned}$$

► **Příklad 9.3.** Dvě firmy vyrábějí telefony, ozn. X_1, resp.X_2 náhodnou veličinu životnost telefonu od první, resp. druhé firmy. Víme, že průměrná životnost telefonů a směrodatná odchylka je 6.5 let a 0.9 let, resp. 5.8 let a 0.8 let v případě první firmy, resp. druhé firmy. Za předpokladu, že náhodně vybereme ze skladu 56 telefonů od první firmy a 64 telefonů od druhé firmy, jaká je pravděpodobnost, že průměrná životnost telefonů od první firmy bude alespoň o rok větší než průměrná životnost telefonů od druhé firmy?
Nápověda: Jaké rozdělení má náhodná veličina $Y = \overline{X_1} - \overline{X_2}$?

► **Příklad 9.4.** Pravděpodobnost, že náhodně oslovený člověk odpoví na anketní otázku je 0.4. Jaká je pravděpodobnost, že oslovíte-li 1600 lidí, anketní otázku s Vámi vyplní více než 650 lidí? [0.296]

► **Příklad 9.5.** Předpokládejte, že měsíční výdaje domácností na kulturu mají $\mathcal{N}(2000, 25000)$.

1. Jaká je pravděpodobnost, že měsíční výdaje jedné domácnosti překročí 2100 Kč? [0.42]
2. Jaká je pravděpodobnost, že průměrné výdaje pěti domácnosti překročí 2100 Kč? [0.08]

► **Příklad 9.6.** Písemná zkouška se skládá z 200 otázek. Student může vybírat jednu ze čtyř odpovědí. Vždy jen jedna je správná. Jaká je pravděpodobnost, že student, který odpovědi vybírá náhodně bez jakékoliv znalosti problematiky, bude mít 25 až 30 odpovědí správně? Zkuste binomické rozdělení aproximovat normálním a porovnejte výsledky.

► **Příklad 9.7.** Nechť X je náhodná veličina s normálním rozdělením. Určete hustotu pro náhodnou veličnu $Y = e^X$. Tato náhodná veličina má tzv. logaritmicko-normální rozdělení. $[g_Y(y) = 1/\sqrt{2\pi y^2 y^{\ln y}}]$

► **Příklad 9.8.** Náhodná veličina X má střední hodnotu $\mu_X = 10$ a rozptyl $\sigma_X^2 = 4$. Určete:

1. $P(\{|X - 10| \ge 3\})$ [nejvýše 4/9]
2. $P(\{|X - 10| < 3\})$ [nejméně 5/9]
3. $P(\{5 < X < 15\})$ [nejméně 21/25]
4. Určete konstantu c tak, aby $P(\{|X - 10| \ge c\}) \le 0.04$ [10]

► **Příklad 9.9.** Životnost svíčky v km má $\mathcal{N}(10\,000, 3000^2)$. Jaká je pravděpodobnost, že při cestě dlouhé 5000 km nebudete muset měnit ani jednu ze čtyř svíček? [0.822]

► **Příklad 9.10.** Chyba(v milimetrech) při měření výšky budovy má $\mathcal{N}(0, 200)$. Kolikrát je třeba měření opakovat, má-li s pravděpodobností 0.9 být alespoň jedna chyba menší než 4 mm? [10]

► **Příklad 9.11.** O prezidentský post se ucházejí tři kandidáti, ozn. A,B,C. Prezident se volí přímo a tak v rámci předvolebního průzkumu jsou náhodně a nezávisle vybraní voliči dotázáni, komu v následujících prezidentských volbách dají hlas. Pro každého kandidáta je určeno předpokládané procento získaných hlasů ve volbách $p_X, X \in \{A, B, C\}$ na základě relativní frekvence hlasů. Kolik lidí musí být osloveno, aby následující tvrzení bylo pravdivé: Pro každého kandidáta X je pravděpodobnost, že předpokládané procento získaných hlasů p_X je správné s odchylkou 1%, minimálně 95%? $[n \ge 10\,000]$

Reference

[1] Capinski M., Zastawniak T. J.: Probability Through Problems
Springer 2001, ISBN 978-0-387-95063-1.

[2] Kerns G. J: Introduction to Probability and Statistics Using R, First Edition
http://cran.r-project.org/web/packages/IPSUR/vignettes/IPSUR.pdf

[3] Kenneth Baclawski: Introduction to Probability with R
Chapman and Hall/CRC, ISBN 978-1420065213.

[4] Walpole R. E, Myers R., Myers S, Ye K. : Probability & Statistics for Engineers & Scientists
Prentice Hall, ISBN:0-13-098469-8