MUNI
ECON

Diskrétní náhodné veličiny a jejich pravděpodobnostní rozdělení

Povinná literatura: Mann (2016, 2024), Kapitola 5

Z čeho studovat pátou lekci?

Povinná literatura:	Mann (2016, 2024), kap. 05
Příprava na cvičení:	Leaflet 06 Koncepty a procedury, cv. 06, kap. 05
Příprava na zkoušku:	Mann (2016, 2024), kap. 05 Leaflet 06 Sbírka úloh, kap. 05 Koncepty a procedury, cv. 06, kap. 05

Motivační vstup - Celostátní loterie

V hypotetické zemi existuje okamžitá loterie s názvem Celostátní loterie, kde každý los stojí 5 dolarů. Hráči mohou vyhrát výhry ve výši 500 000, 10 000, 1 000, 100, 50, 20 a 10 dolarů. Každý los má sedm políček – horní políčko Výhra a šest Hráčových políček. Hráč vyhraje, pokud některé z jeho políček obsahuje částku shodující se s výherní částkou v horním políčku. V opačném případě prohrává.

Výhra (v dolarech)	Počet losů	Výhra (v dolarech)	Počet losů
0	18 000 000	100	5 000
10	1 620 000	1 000	730
20	364 000	10 000	200
50	10 000	500 000	70

Pokud si koupíte 1 los, jaký očekáváte zisk?

Náhodná veličina

Náhodná veličina je proměnná, jejíž hodnota je určena výsledkem náhodného experimentu.

Proměnná, která nabývá spočetné množství hodnot, se nazývá **diskrétní náhodná veličina**.

Proměnná, která může nabývat jakékoli hodnoty obsažené v jednom nebo více intervalech, se nazývá **spojitá náhodná veličina**.

Náhodná veličina - Příklady

Diskrétní náhodná veličina.

1. Počet prodaných aut v autosalonu během určitého měsíce
2. Počet domů v určitém bloku
3. Počet stížností přijatých u letecké společnosti v určitý den
4. Počet zákazníků, kteří navštíví banku během libovolné hodiny
5. Počet padlých hlav při třech hodech mincí

Spojitá náhodná veličina.

1. Délka místnosti
2. Doba cesty z domova do práce
3. Množství mléka v cisterně
4. Hmotnost dopisu
5. Cena domu

Pravděpodobnostní rozdělení diskrétní náhodné veličiny

Pravděpodobnostní rozdělení diskrétní náhodné veličiny uvádí všechny možné hodnoty, které může náhodná veličina nabývat, a jejich odpovídající pravděpodobnosti.

Dvě vlastnosti pravděpodobnostního rozdělení

Pravděpodobnostní rozdělení diskrétní náhodné proměnné má dvě následující vlastnosti.

1. $0 \leq P(x) \leq 1$ pro každou hodnotu x
2. $\sum P(x) = 1$

Příklad 1

Zadání: Mějme tabulkou zadánu četnost a relativní četnost počtu vozidel vlastněných rodinami. Nechť x je počet vozidel vlastněných náhodně vybranou rodinou. Napište pravděpodobnostní rozdělení x a vytvořte histogram pro toto pravděpodobnostní rozdělení.

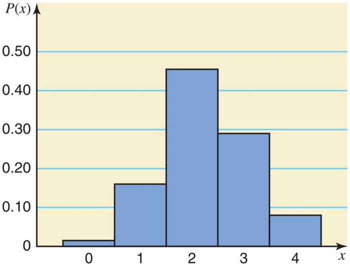
Number of Vehicles Owned	Frequency	Relative Frequency
0	30	0.015
1	320	0.160
2	910	0.455
3	580	0.290
4	160	0.080
	$N = 2000$	**Sum = 1.000**

Příklad 1: Řešení

Obrázek: Rozdělení

Number of Vehicles Owned x	Probability $P(x)$
0	0.015
1	0.160
2	0.455
3	0.290
4	0.080
	Sum = 1.000

Obrázek: Histogram



Příklad 2

Zadání: Pomocí pravděpodobnostního rozdělení uvedeného v tabulce najděte pravděpodobnost, že

(a) náhodně vybraná rodina vlastní dvě vozidla.

(b) náhodně vybraná rodina vlastní alespoň dvě vozidla.

(c) náhodně vybraná rodina vlastní nejvýše jedno vozidlo.

(d) náhodně vybraná rodina vlastní tři nebo více vozidel.

Number of Vehicles Owned x	**Probability** $P(x)$
0	0.015
1	0.160
2	0.455
3	0.290
4	0.080
	Sum = 1.000

Příklad 2: Řešení

a) Pravděpodobnost, že náhodně vybraná rodina vlastní dvě vozidla, je $P(2) = 0.455$

b) Pravděpodobnost, že náhodně vybraná rodina vlastní alespoň dvě vozidla, je $P(2 \text{ nebo } 3 \text{ nebo } 4) = P(2) + P(3) + P(4) = 0.455 + 0.290 + 0.080 = 0.825$

c) Pravděpodobnost, že náhodně vybraná rodina vlastní nejvýše jedno vozidlo, je
$P(0 \text{ nebo } 1) = P(0) + P(1) = 0.015 + 0.160 = 0.175$

d) Pravděpodobnost, že náhodně vybraná rodina vlastní tři nebo více vozidel, je
$P(3 \text{ nebo } 4) = P(3) + P(4) = 0.290 + 0.080 = 0.370$

Příklad 3

Zadání: Každá z následujících tabulek uvádí určité hodnoty x a jejich příslušné pravděpodobnosti. Určete, zda každá tabulka představuje platné pravděpodobnostní rozdělení.

(a)

x	$P(x)$
0	.08
1	.11
2	.39
3	.27

(b)

x	$P(x)$
2	.25
3	.34
4	.28
5	.13

(c)

x	$P(x)$
7	.70
8	.50
9	−.20

Řešení:

(a) Ne, protože součet všech pravděpodobností není roven 1.

(b) Ano.

(c) Ne, protože jedna z pravděpodobností je záporná.

Střední hodnota a směrodatná odchylka diskrétní náhodné veličiny

Střední hodnota diskrétní náhodné veličiny x je hodnota, která se v průměru očekává při každém jednom opakování experimentu, pokud je tento experiment opakován velký počet krát. Je označena μ a vypočítává se jako

$$\mu = \sum xP(x).$$

Střední hodnota diskrétní náhodné proměnné x je také nazývána její očekávaná hodnota a je označena $E(x)$; to znamená,

$$E(x) = \sum xP(x).$$

Střední hodnota a směrodatná odchylka diskrétní náhodné veličiny

Směrodatná odchylka diskrétní náhodné proměnné x měří variabilitu jejího pravděpodobnostního rozdělení a vypočítává se jako

$$\sigma = \sqrt{\sum x^2 P(x) - \mu^2}.$$

Poznámka: Směrodatná odchylka diskrétní náhodné veličiny může být interpretována nebo používána stejným způsobem jako směrodatná odchylka datového souboru v sekci 3.4 kapitoly 3.

Příklad 4

Zadání: Společnost Věda&Krása plánuje uvést na trh nový kosmetický výrobek. Podle analýzy finančního oddělení společnosti, pokud tento produkt dosáhne vysokých prodejů, vydělá ročně 4.5 milionu dolarů, pokud budou prodeje průměrné, vydělá ročně 1.2 milionu dolarů a pokud budou prodeje nízké přijde o 2.3 milionu dolarů ročně. Pravděpodobnosti těchto tří scénářů jsou postupně 0.32, 0.51 a 0.17.

a) Nechť x označuje zisk (v milionech dolarů) vydělaný společností z tohoto produktu za rok. Zapište pravděpodobnostní rozdělení x.

b) Vypočítejte střední hodnotu a směrodatnou odchylku x.

Příklad 4: Řešení a)

Tabulka níže uvádí pravděpodobnostní rozdělení x. Poznamenejme, že protože x označuje zisk vydělaný společností, je ztráta v tabulce zapsána jako negativní zisk.

x	$P(x)$
4.5	.32
1.2	.51
−2.3	.17

Příklad 4: Řešení b)

(b) Tabulka ukazuje všechny údaje potřebné pro výpočet střední hodnoty a směrodatné odchylky x.

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
4.5	.32	1.440	20.25	6.4800
1.2	.51	.612	1.44	.7344
−2.3	.17	−.391	5.29	.8993
		$\Sigma xP(x) = 1.661$		$\Sigma x^2P(x) = 8.1137$

$$\mu = \sum xP(x) = 1.661 \text{ milionů dolarů}$$

$$\sigma = \sqrt{\sum x^2 P(x) - \mu^2} = \sqrt{8.1137 - (1.661)^2} = 2.314 \text{ milionů dolarů}$$

Motivační vstup - Celostátní loterie

V hypotetické zemi existuje okamžitá loterie s názvem Celostátní loterie, kde každý los stojí 5 dolarů. Hráči mohou vyhrát výhry ve výši 500 000, 10 000, 1 000, 100, 50, 20 a 10 dolarů. Každý los má sedm políček – horní políčko Výhra a šest Hráčových políček. Hráč vyhraje, pokud některé z jeho políček obsahuje částku shodující se s výherní částkou v horním políčku. V opačném případě prohrává.

Výhra (v dolarech)	Počet losů	Výhra (v dolarech)	Počet losů
0	18 000 000	100	5 000
10	1 620 000	1 000	730
20	364 000	10 000	200
50	10 000	500 000	70

Pokud si koupíte 1 los, jaký očekáváte zisk?

Celostátní loterie

Čistá výhra hráče pro každý výherní tiket je rovna výši výhry minus 5 dolarů, což je cena tiketu. Čistý zisk pro každý nevýherní tiket je tedy -5 dolarů, což je cena tiketu. Označme:

$$x = \text{čistá částka, kterou hráč vyhraje při hraní této loterie}$$

Tabulka na následujícím slidu ukazuje pravděpodobnostní rozdělení x a všechny výpočty potřebné k určení střední hodnoty x pro toto pravděpodobnostní rozdělení. Pravděpodobnost výsledku (čisté výhry) se vypočítá tak, že se počet tiketů s daným výsledkem (výhrou) vydělí celkovým počtem tiketů.

Celostátní loterie

x (dolarů)	P(x)	xP(x)
-5	18 000 000/20 000 000 = 0.9000000	-4.5000000
5	1 620 000/20 000 000 = 0.0810000	0.4050000
15	364 000/20 000 000 = 0.0182000	0.2730000
45	10 000/20 000 000 = 0.0005000	0.0225000
95	5 000/20 000 000 = 0.0002500	0.0237500
995	730/20 000 000 = 0.0000365	0.0363175
9995	200/20 000 000 = 0.0000100	0.0999500
499 995	70/20 000 000 = 0.0000035	1.7499825
$\sum xP(x)$	**-1.8895000**	

Celostátní loterie

Proto je střední hodnota nebo očekávaná hodnota x rovna:

$$\mu = \sum xP(x) = -1.8895000 \approx -1.89\,\text{dolarů}.$$

Tato hodnota představuje očekávanou hodnotu náhodné proměnné x, tedy $E(X) = \sum xP(x) = -1.89\,\text{dolarů}$.

Střední hodnota čistých výher z loterie je -1.89 dolarů. Jinými slovy, všichni hráči dohromady ztratí v průměru 1.89 dolaru na každý tiket. To znamená, že z každých 5 dolarů (cena tiketu) se 3.11 dolaru vrátí hráčům ve formě výher a 1.89 dolaru půjde státu, který pokryje náklady na provoz loterie, provize vyplacené agentům a zisk pro stát.

Všimněte si, že 1.89 dolaru představuje 37.8 % z 5 dolarů. Můžeme tedy také uvést, že 37.8 % z celkové částky, kterou hráči utratí za tuto loterii, půjde státu a $100 - 37.8 = 62.2\%$ bude vráceno hráčům ve formě výher.

Binomický experiment

Binomický experiment musí splňovat čtyři následující podmínky.

1. Existuje n identických pokusů.
2. Každý pokus má pouze dva možné výsledky (nebo jevy).
3. Pravděpodobnosti obou výsledků (nebo jevů) zůstávají pro každý pokus neměnné.
4. Pokusy jsou nezávislé.

Příklad 5

Zadání: Zvažte experiment, který se skládá z 10 hodů mincí. Určete, zda se jedná o binomický experiment.

Řešení:

- Existuje celkem $n = 10$ pokusů (hodů) a všechny jsou identické.
- Každý pokus (hod) má pouze dva možné výsledky: panna a orel.
- Pravděpodobnost získání hlavy (úspěchu) je $\frac{1}{2}$ a orla (neúspěchu) je $\frac{1}{2}$ pro jakýkoliv hod. Tzn., $p = P(H) = \frac{1}{2}$ a $q = P(O) = \frac{1}{2}$.
- Pokusy (hody) jsou nezávislé.

Experiment skládající se z 10 hodů je tedy binomický experiment.

Příklad 6

Zadání:

(a) Sedmdesát pět procent studentů na vysoké škole (s velkým počtem studentů) používá Instagram. Je vybrán vzorek pěti studentů z této vysoké školy a ti jsou dotázáni, zda Instagram používají. Je tento experiment binomickým experimentem?

(b) Ve skupině 12 studentů na vysoké škole jich 9 používá Instagram. Je vybráno pět studentů z této skupiny a ti dotázáni, zda Instagram používají. Je tento experiment binomickým experimentem?

Příklad 6: Řešení (a)

Zkontrolujeme, zda jsou splněny všechny čtyři podmínky binomického pravděpodobnostního rozdělení.

1. Tento příklad se skládá z pěti identických pokusů.
2. Každý pokus má dva možné výsledky: student používá Instagram nebo student nepoužívá Instagram.
3. Pravděpodobnost p, že student používá Instagram je .75. Pravděpodobnost q, že student nepoužívá Instagram je .25.
4. Každý pokus (student) je nezávislý.

Protože jsou splněny všechny čtyři podmínky binomického experimentu, jedná se o příklad binomického experimentu.

Příklad 6: Řešení (b)

Zkontrolujeme, zda jsou splněny všechny čtyři podmínky binomického pravděpodobnostního rozdělení.

1. Tento příklad se skládá z pěti identických pokusů.
2. Každý pokus má dvě možné výsledky: student používá Instagram nebo student nepoužívá Instagram.
3. Pravděpodobnost p je, že student používá Instagram. Pravděpodobnost q je, že student nepoužívá Instagram. Tyto pravděpodobnosti nezůstávají konstantní pro každý výběr. Pravděpodobnost každého výsledku se mění s každým výběrem v závislosti na tom, co se stalo v předchozích výběrech.
4. Protože pravděpodobnosti p a q nezůstávají stálé pro každý výběr, pokusy nejsou nezávislé.

Vzhledem k tomu, že třetí a čtvrtá podmínka binomického experimentu není splněna, nejedná se o příklad binomického experimentu.

Alternativní (Bernoulliho) rozdělení

- Alternativní (Bernoulliho) rozdělení je diskrétní rozdělení pravděpodobnosti náhodné veličiny, která nabývá hodnoty 1 s pravděpodobností p a hodnoty 0 s pravděpodobností $1-p$. Jde o speciální případ binomického rozdělení s $n=1$.

Pravděpodobnostní funkce:

$$P(x) = p^x(1-p)^{1-x}, \quad x \in \{0,1\},$$

kde

- p – pravděpodobnost úspěchu (hodnota 1)
- x – realizace náhodné veličiny, kde $x \in \{0,1\}$

Charakteristiky:

Střední hodnota: $\mu = p$ Směrodatná odchylka: $\sigma = \sqrt{p(1-p)}$

Binomické rozdělení a binomický vzorec

Pro binomický experiment je pravděpodobnost přesně x úspěchů z n pokusů dána binomickým vzorcem

$$P(x) = C_x^n p^x q^{n-x}$$

kde

$$\begin{aligned} n &= \text{celkový počet pokusů} \\ p &= \text{pravděpodobnost úspěchu} \\ q = 1 - p &= \text{pravděpodobnost neúspěchu} \\ x &= \text{počet úspěchů v } n \text{ pokusech} \\ n - x &= \text{počet neúspěchů v } n \text{ pokusech} \end{aligned}$$

Příklad 7

Zadání: Sedmdesát pět procent studentů vysoké školy s velkým počtem studentů používá sociální síť Instagram. Tři studenti z této vysoké školy jsou náhodně vybráni. Jaká je pravděpodobnost, že přesně dva z těchto tří studentů používají Instagram?

Řešení: Ze zadání víme, že: $n = 3$, $x = 2$, a $p = 0.75$.
Pravděpodobnost dvou úspěchů je označena $P(x = 2)$ nebo $P(2)$.

Number of ways to obtain 2 success in 3 trials

Number of successes

Number of failures

$$P(2) = {}_3C_2\,(.75)^2\,(.25)^1 = (3)(.5625)\,(.25) = \mathbf{.4219}$$

Probability of success

Probability of failure

Příklad 8

Zadání: Společnost Express House Delivery Service klade důraz na poskytování kvalitních služeb a garantuje vrácení všech poplatků, pokud balík není doručen včas. Z minulých dat vyplývá, že 2 % balíků nejsou doručeny včas. Předpokládejme, že firma v daný den odesílá 10 balíků.

(a) Najděte pravděpodobnost, že přesně jeden z těchto 10 balíků nedorazí do svého určení v stanoveném čase.

(b) Najděte pravděpodobnost, že nejvýše jeden z těchto 10 balíků nedorazí do svého určení v stanoveném čase.

Příklad 8: Řešení (a)

Ze zadání víme:

$$\begin{aligned} n &= \text{celkový počet odeslaných balíků} = 10 \\ p &= P(\text{úspěch}) = 0.02 \\ q &= P(\text{neúspěch}) = 1 - 0.02 = 0.98 \\ x &= \text{počet úspěchů} = 1 \\ n - x &= \text{počet neúspěchů} = 10 - 1 = 9 \end{aligned}$$

Řešení: (a)

$$\begin{aligned} P(x = 1) &= C_1^{10}(.02)^1(.98)^9 = \frac{10!}{1!(10-1)!}(.02)^1(.98)^9 \\ &= (10)(.02)(.83374776) = .1667 \end{aligned}$$

Existuje .1667 pravděpodobnost, že přesně jeden z 10 odeslaných balíků nedorazí do svého určení ve stanoveném čase.

Příklad 8: Řešení (b)

Řešení: (b) Pravděpodobnost, že nejvýše jeden ze deseti balíků nebude doručen, je dána součtem pravděpodobností $x = 0$ a $x = 1$

$$\begin{aligned} P(x \leq 1) &= P(x = 0) + P(x = 1) \\ &= C_0^{10}(.02)^0(.98)^{10} + C_1^{10}(.02)^1(.98)^9 \\ &= (1)(1)(.81707281) + (10)(.02)(.83374776) \\ &= .8171 + .1667 = .9838 \end{aligned}$$

Tedy pravděpodobnost, že nejvýše jeden z deseti odeslaných balíků nedorazí do svého určení v daném čase, je .9838.

Příklad 9

Zadání: Podle průzkumu neplánuje 33 % amerických zaměstnanců v blízké budoucnosti změnit své pracovní místo. Nechť x označuje počet zaměstnanců v náhodném vzorku tří amerických zaměstnanců, kteří neplánují v blízké budoucnosti změnit své pracovní místo. Napište pravděpodobnostní rozdělení x a nakreslete histogram pro toto pravděpodobnostní rozdělení.

$n = \text{celkový počet zaměstnanců ve vzorku} = 3$

$p = P(\text{zaměstnanec neplánuje v blízké budoucnosti změnit práci}) = .33$

$q = P(\text{zaměstnanec plánuje v blízké budoucnosti změnit práci}) = 1 - .33$

Příklad 9: Řešení

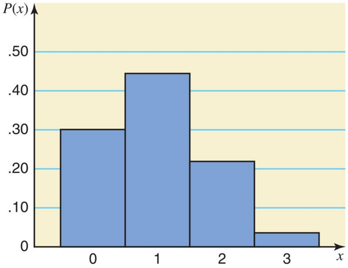
$$P(0) = C_0^3(.33)^0(.67)^3 = (1)(1)(.300763) = .3008$$
$$P(1) = C_1^3(.33)^1(.67)^2 = (3)(.33)(.4489) = .4444$$
$$P(2) = C_2^3(.33)^2(.67)^1 = (3)(.1089)(.67) = .2189$$
$$P(3) = C_3^3(.33)^3(.67)^0 = (1)(.035937)(1) = .0359$$

Obrázek: Rozdělení

Table 5.9 Probability Distribution of x

x	$P(x)$
0	.3008
1	.4444
2	.2189
3	.0359

Obrázek: Histogram



Použití tabulky binomických pravděpodobností

V rámci programu Excel lze k získání pravděpodobnosti využít příkaz

$$= BINOM.DIST(x, n, p, cumul)$$

kde:

- x : Počet úspěchů, pro které chcete vypočítat pravděpodobnost.
- n : Počet pokusů.
- p : Pravděpodobnost úspěchu v jednom pokusu.
- $cumul$: Logická hodnota. Pokud je *TRUE*, vrátí kumulativní pravděpodobnost, že dojde k nanejvýš x úspěchům. Pokud je *FALSE*, vrátí pravděpodobnost, že dojde přesně k x úspěchům.

Použití tabulky binomických pravděpodobností

Pravděpodobnosti pro binomický experiment lze také vyčíst z tabulky I v příloze B učebnice Mann, popřípadě z jiných zdrojů kde jsou vypsány tabulky binomických pravděpodobností.

Tato tabulka uvádí pravděpodobnosti x pro n od 1 do 25.

Tato tabulka uvádí pravděpodobnosti x pro vybrané hodnoty p.

Příklad 10

Zadání: Podle průzkumu 30 % vysokoškolských studentů uvedlo, že tráví příliš mnoho času na Facebooku. Předpokládejme, že to platí pro celou populaci studentů. Je vybrán náhodný vzorek šesti studentů. Pomocí tabulky I v příloze B zodpovězte následující otázky:

Příklad 10

Zadání:

(a) Určete pravděpodobnost, že přesně tři z těchto šesti studentů VŠ uvedou, že tráví na Facebooku příliš mnoho času.

(b) Určete pravděpodobnost, že nejvýše dva z těchto šesti studentů VŠ uvedou, že tráví na Facebooku příliš mnoho času.

(c) Určete pravděpodobnost, že alespoň tři z těchto šesti studentů VŠ uvedou, že tráví na Facebooku příliš mnoho času.

(d) Určete pravděpodobnost, že jeden až tři z těchto šesti studentů VŠ řeknou, že tráví příliš mnoho času na Facebooku.

(e) Ať x označuje počet osob z náhodného vzorku šesti studentů VŠ, kteří řeknou, že tráví příliš mnoho času na Facebooku. Napište pravděpodobnostní rozdělení x a nakreslete histogram pro toto pravděpodobnostní rozdělení.

Příklad 10

Tabulka rozdělení: Příklad nalezení $P(x = 3)$ pro $n = 6$ a $p = 0.30$

$p = .30$

		p					
n	**x**	**.05**	**.10**	**.20**	**.30**	**…**	**.95**
$n = 6 \longrightarrow$ 6	0	.7351	.5314	.2621	.1176	…	.0000
	1	.2321	.3543	.3932	.3025	…	.0000
	2	.0305	.0984	.2458	.3241	…	.0001
	$x = 3 \longrightarrow$ 3	.0021	.0146	.0819	.1852	…	.0021
	4	.0001	.0012	.0154	.0595	…	.0305
	5	.0000	.0001	.0015	.0102	…	.2321
	6	.0000	.0000	.0001	.0007	…	.7351

$P(x = 3) = .1852$

Příklad 10: Řešení

Tabulka rozdělení: Část tabulky I pro $n = 6$ a $p = 0.30$

		p
n	x	**.30**
6	0	.1176
	1	.3025
	2	.3241
	3	.1852
	4	.0595
	5	.0102
	6	.0007

Příklad 10: Řešení

$$
\begin{aligned}
&\textbf{(a) } P(3) &&= .1852 \\
&\textbf{(b) } P(\text{nejvýše dva } 2) &&= P(0 \text{ nebo } 1 \text{ nebo } 2) \\
& &&= P(0) + P(1) + P(2) \\
& &&= .1176 + .3025 + .3241 = .7442 \\
&\textbf{(c) } P(\text{alespoň tři } 3) &&= P(3 \text{ nebo } 4 \text{ nebo } 5 \text{ nebo } 6) \\
& &&= P(3) + P(4) + P(5) + P(6) \\
& &&= .1852 + .0595 + .0102 + .0007 = .2556 \\
&\textbf{(d) } P(1 \text{ až } 3) &&= P(1) + P(2) + P(3) \\
& &&= .3025 + .3241 + .1852 = .8118
\end{aligned}
$$

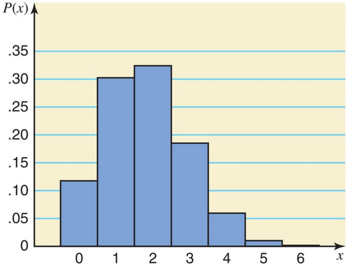
Příklad 10: Řešení (e)

Rozdělení pravděpodobnosti: náhodné veličiny x pro $n = 6$ a $p = 0.30$

Obrázek: Rozdělení

x	$P(x)$
0	.1176
1	.3025
2	.3241
3	.1852
4	.0595
5	.0102
6	.0007

Obrázek: Histogram



Pravděpodobnost úspěchu a tvar binomického rozdělení

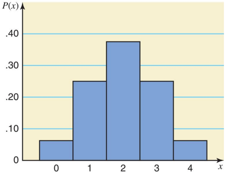
Pro libovolný počet pokusů n:

1. Binomické pravděpodobnostní rozdělení je symetrické, pokud $p = .50$.
2. Binomické pravděpodobnostní rozdělení je zešikmené doprava, pokud p je menší než .50.
3. Binomické pravděpodobnostní rozdělení je zešikmené doleva, pokud p je větší než .50.

Tvar binomického rozdělení náhodné veličiny x

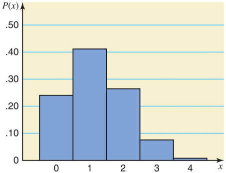
pro
$n = 4$
$p = 0.50$

x	$P(x)$
0	.0625
1	.2500
2	.3750
3	.2500
4	.0625



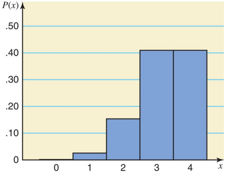
pro
$n = 4$
$p = 0.30$

x	$P(x)$
0	.2401
1	.4116
2	.2646
3	.0756
4	.0081



pro
$n = 6$
$p = 0.80$

x	$P(x)$
0	.0016
1	.0256
2	.1536
3	.4096
4	.4096



Střední hodnota a směrodatná odchylka binomického rozdělení

Střední hodnota a směrodatná odchylka binomického rozdělení lze vypočítat jako:

$$\mu = np$$

a

$$\sigma = \sqrt{npq}$$

kde n je celkový počet pokusů, p je pravděpodobnost úspěchu a q je pravděpodobnost neúspěchu.

Příklad 11

Zadání:

Podle průzkumu Pew Research Center z května 2015 se 22.8 % dospělých Američanů nehlásí k žádnému náboženství.

Předpokládejme, že zjištění platí pro celou populaci dospělých.

Je vybrán náhodný vzorek 50 dospělých Američanů. Označme x jako počet těch, kteří se nehlásí k žádnému náboženství. Určete střední hodnotu a směrodatnou odchylku pravděpodobnostního rozdělení x.

Příklad 11: Řešení

Ze zadání známe: $n = 50, p = .228,$ a $q = .772.$
Použitím vzorců pro střední hodnotu a směrodatnou odchylku binomického rozdělení získáme:

$$\mu = np = 50(0.228) = 11.4$$
$$\sigma = \sqrt{npq} = \sqrt{(50)(0.228)(0.772)} = 2.9666$$

Hypergeometrické rozdělení pravděpodobnosti

Pravděpodobnost x, že při výběru n prvků z množiny o velikosti N, v níž má r prvků požadovanou vlastnost, bude mít právě x prvků tuto vlastnost.

$$P(x) = \frac{C_x^r \cdot C_{n-x}^{N-r}}{C_n^N}$$

kde

$$\begin{aligned}
N &= \text{celkový počet prvků v populaci} \\
r &= \text{počet úspěchů v populaci} \\
N - r &= \text{počet neúspěchů v populaci} \\
n &= \text{počet pokusů (velikost vzorku)} \\
x &= \text{počet úspěchů v } n \text{ pokusech} \\
n - x &= \text{počet neúspěchů v } n \text{ pokusech}
\end{aligned}$$

Použití tabulky binomických pravděpodobností

V rámci programu Excel lze k získání pravděpodobnosti využít příkaz

$$= HYPGEOM.DIST(x, n, K, N, cumul)$$

kde:

- x : Počet úspěchů ve výběru (počet požadovaných úspěchů).
- n : Počet položek ve výběru (velikost vzorku).
- K : Počet úspěchů v populaci.
- N : Velikost celé populace..
- $cumul$: Logická hodnota. Pokud je *TRUE*, vrátí kumulativní pravděpodobnost, že dojde k nanejvýš x úspěchům. Pokud je *FALSE*, vrátí pravděpodobnost, že dojde přesně k x úspěchům.

Příklad 12

Zadání: Společnost Brown Manufacturing vyrábí autodíly, které se prodávají auto dealerům. Minulý týden společnost dodala dealerovi 25 autodílů. Později zjistila, že 5 z těchto dílů bylo vadných. Do té doby, než manažer společnosti kontaktoval dealera, byly již 4 autodíly z této zásilky prodány. Jaká je pravděpodobnost, že 3 z těch 4 dílů byly dobré a 1 byl vadný?

Příklad 12: Řešení

Ze zadání známe:

$N = 25, \quad r = 20, \quad N - r = 5, \quad n = 4, \quad x = 3, \quad n - x = 1$

$$P(x) = \frac{C_x^r \cdot C_{n-x}^{N-r}}{C_n^N} = \frac{C_3^{20} \cdot C_1^5}{C_4^{25}} = \frac{\frac{20!}{3!(20-3)!} \cdot \frac{5!}{1!(5-1)!}}{\frac{25!}{4!(25-4)!}} = \frac{(1140)(5)}{12\,650} = .4506$$

Takže pravděpodobnost, že tři ze čtyř prodaných dílů jsou dobré a jeden je vadný, je .4506.

Příklad 13

Zadání: Společnost Dawn Corporation má 12 zaměstnanců na manažerských pozicích. Z toho 7 je žen a 5 mužů. Společnost plánuje poslat 3 z těchto 12 manažerů na konferenci. Určete pravděpodobnost, pokud jsou náhodně vybráni 3 manažeři z 12,

(a) že všichni 3 jsou ženy.

(b) že maximálně 1 z nich je žena.

Příklad 13: Řešení (a)

Ze zadání známe:

$N = 12, \quad r = 7, \quad N - r = 5, \quad n = 3, \quad x = 3, \quad n - x = 0$

$$P(3) = \frac{C_x^r \cdot C_{n-x}^{N-r}}{C_n^N} = \frac{C_3^7 \cdot C_0^5}{C_3^{12}} = \frac{(35)(1)}{220} = .1591$$

Pravděpodobnost, že všechny 3 vybraní manažeři jsou ženy je .1591.

Příklad 13: Řešení (b)

Ze zadání známe:

$N = 12, \quad r = 7, \quad N - r = 5, \quad n = 3, \quad x = 0 \text{ a } 1, \quad n - x = 3 \text{ a } 2$

$$
\begin{aligned}
P(0) &= \frac{C_x^r \cdot C_{n-x}^{N-r}}{C_n^N} = \frac{C_0^7 \cdot C_3^5}{C_3^{12}} = \frac{(1)(10)}{220} &&= 0.0455 \\
P(1) &= \frac{C_x^r \cdot C_{n-x}^{N-r}}{C_n^N} = \frac{C_1^7 \cdot C_2^5}{C_3^{12}} = \frac{(7)(10)}{220} &&= 0.3182 \\
P(x \leq 1) &= P(0) + P(1) = 0.0455 + 0.3182 &&= 0.3637
\end{aligned}
$$

Pravděpodobnost, že maximálně 1 ze 3 vybraných manažerů je žena, je .3637.

Poissonovo rozdělení pravděpodobnosti

Pro použití Poissonova rozdělení musí být splněny tři podmínky.

1. x je diskrétní náhodná veličina.
2. Výskyty jsou náhodné.
3. Výskyty jsou nezávislé.

Příklady Poissonova rozdělení pravděpodobnosti

1. Počet nehod, ke kterým dojde na určité dálnici během jednoho týdne.
2. Počet zákazníků vstupujících do potravin během jednoho hodinového intervalu.
3. Počet prodaných televizorů v obchodním domě během jednoho týdne.

Vzorec Poissonova rozdělení pravděpodobnosti

Podle **Poissonova rozdělení pravděpodobnosti** je pravděpodobnost x výskytů na daném intervalu dána jako

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

kde λ (vyslovuje se lambda) je průměrný počet výskytů v daném intervalu a hodnota e je přibližně 2.71828.

Příklad 14

Zadání: Domácnost dostává průměrně 9.5 telefonních hovorů od telemarketingu za týden. Pomocí vzorce Poissonova rozdělení určete pravděpodobnost, že náhodně vybraná domácnost obdrží přesně 6 telemarketingových hovorů během daného týdne.

Řešení:

$$\begin{aligned} P(6) &= \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(9.5)^6 e^{-9.5}}{6!} \\ &= \frac{(735091.8906)(.00007485)}{720} \\ &= 0.0764 \end{aligned}$$

Příklad 15

Zadání: Pračka v prádelně se průměrně pokazí třikrát za měsíc. Použitím vzorce Poissonova rozdělení pravděpodobnosti naleznete pravděpodobnost, že během příštího měsíce bude mít tato pračka

(a) přesně dvě poruchy

(b) nejvýše jednu poruchu

Příklad 15

(a) P(přesně dvě poruchy)

$$
\begin{aligned}
P(x=2) &= \frac{\lambda^2 e^{-\lambda}}{2!} = \frac{3^2 e^{-3}}{2} \\
&= 9 \cdot 0.04978707 \\
&= 0.2240
\end{aligned}
$$

(b) P(nejvýše jedna porucha)

$$
\begin{aligned}
P(0 \text{ nebo } 1 \text{ porucha}) &= P(x=0) + P(x=1) \\
&= \frac{3^0 e^{-3}}{0!} + \frac{3^1 e^{-3}}{1!} \\
&= 1 \cdot 0.04978707 + 3 \cdot 0.04978707 \\
&= 0.0498 + 0.1494 \\
&= 0.1992
\end{aligned}
$$

Příklad 16

Zadání: Martin poskytuje 7denní bezplatné vyzkoušení produktů s možností vrácení a plnou náhradou, pokud zákazník není spokojen. Na základě minulých dat jsou v průměru 2 z 10 prodaných produktů vráceny. Pomocí Poissonova rozdělení vypočítejte pravděpodobnost, že přesně 6 z 40 prodaných produktů v daný den bude vráceno a náhrada bude vyplacena.

Řešení: $\lambda = 8, \quad x = 6$

$$P(6) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{8^6 e^{-8}}{6!} = \frac{(262\,144)(0.00033546)}{720} = 0.1221$$

Pravděpodobnost, že přesně 6 produktů ze 40 prodaných v daný den bude vráceno, je 0.1221.

Tabulky pravděpodobností Poissonova rozdělení

V rámci programu Excel lze k získání pravděpodobnosti využít příkaz

$$= POISSON.DIST(x, mean, cumul)$$

kde:

- *x* : Počet událostí, pro které chcete vypočítat pravděpodobnost.
- *mean* : Očekávaný počet událostí na intervalu (střední hodnota).
- *cumul* : Logická hodnota. Pokud je *TRUE*, vrátí kumulativní pravděpodobnost, že dojde k nanejvýš x událostem. Pokud je *FALSE*, vrátí pravděpodobnost, že dojde přesně k x událostem.

Pravděpodobnosti pro Poissonovo rozdělení mohou být také přečteny z tabulky III v příloze B, tabulky Poissonových pravděpodobností.

Příklad 17

V průměru jsou v pobočce komerční banky otevřeny dva nové účty denně. Použitím tabulky III v příloze B najděte pravděpodobnost, že v daný den bude počet nově otevřených účtů v této bance:

(a) přesně 6
(b) nejvýše 3
(c) alespoň 7

Příklad 17: Řešení

			λ		
x	**1.1**	**1.2**	$\cdots$	$\boxed{\mathbf{2.0}}$	$\longleftarrow \lambda = 2.0$
0				.1353	
1				.2707	
2				.2707	
3				.1804	
4				.0902	
5				.0361	
$x = 6 \longrightarrow \boxed{6}$				$\boxed{.0120}$	$\longleftarrow P(x = 6)$
7				.0034	
8				.0009	
9				.0002	

Příklad 17: Řešení

(a) $P(6) = 0.0120$

(b)
$$
\begin{aligned}
P(\text{nejvýše } 3) &= P(0) + P(1) + P(2) + P(3) \\
&= 0.1353 + 0.2707 + 0.2707 + 0.1804 \\
&= 0.8571
\end{aligned}
$$

(c)
$$
\begin{aligned}
P(\text{alespoň } 7) &= P(7) + P(8) + P(9) \\
&= 0.0034 + 0.0009 + 0.0002 \\
&= 0.0045
\end{aligned}
$$

Střední hodnota a směrodatná odchylka Poissonova rozdělení

Střední hodnota a směrodatná odchylka Poissonova rozdělení pravděpodobnosti lze vypočítat jako:

$$\mu = \lambda$$

a

$$\sigma^2 = \lambda$$

$$\sigma = \sqrt{\lambda}$$

Příklad 18

Zadání: Prodavač aut prodává průměrně 0.9 auta za den. Označme x počet aut prodaných tímto prodejcem v daný den. Najděte střední hodnotu, rozptyl a směrodatnou odchylku.

Řešení:

$$\mu = \lambda = 0.9 \text{ aut}$$
$$\sigma^2 = \lambda = 0.9$$
$$\sigma = \sqrt{\lambda} = \sqrt{0.9} \approx 0.949 \text{ aut}$$

Shrnutí přednášky:

Náhodná veličina (NV)

Pravděpodobnostní rozdělení diskrétní NV

Střední hodnota a směrodatná odchylka DNV

Binomické rozdělení pravděpodobnosti

- Binomický experiment
- Binomické rozdělení a binomický vzorec
- Použití tabulky binomických pravděpodobností
- Pravděpodobnost úspěchu a tvar binomického rozdělení
- Střední hodnota a směrodatná odchylka BR

Hypergeometrické rozdělení

Poissonovo rozdělení

- Tabulka pravděpodobností Poissonova rozdělení
- Střední hodnota a směrodatná odchylka PR

Co si nastudovat na následující týden?

Příprava na cvičení: Leaflet 06
Koncepty a procedury, cv. 06, kap. 05

Povinná literatura: Mann (2016, 2024), Kapitola 6

Děkuji za pozornost!

Source: 'Alea Acta Est' by Enrico Chavez

Rozšiřující terminologie (nebude testováno)

V rámci studia statistické literatury se můžete setkat také s pojmem **distribuční funkce**.

Distribuční funkce (funkce rozdělení) udává pravděpodobnost, že hodnota náhodné proměnné je menší než zadaná hodnota. Často se označuje jako (zleva) kumulovaná pravděpodobnost. V angličtině se používá termín Cumulative Distribution Function (CDF).

Detailněji v Hendl (2015)[1] str. 137.

[1] HENDL, Jan. Přehled statistických metod: analýza a metaanalýza dat. Páté, rozšířené vydání. Praha: Portál, 2015. ISBN 978-80-262-0981-2.