Jihočeská univerzita v Českých Budějovicích

Přírodovědecká fakulta

Aplikace matematických znalostí při výuce biologie

Diplomová práce

Bc. Lucie Studená

Školitelka: RNDr. Ing. Jana Kalová, Ph.D.

České Budějovice 2018

Bibliografické údaje

Studená, L., 2018: Aplikace matematických znalostí při výuce biologie [Applications of mathematical knowledge in teaching biology. Mgr. Thesis, in Czech.] – p. 98, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Anotace

The Theses deals with applications of mathematical knowledge in teaching biology and it is divided into four chapters. Each chapter is dedicated to another application: 1. Application of conditional probability in medical diagnostics, 2. Application of exponential function in population ecology, 3. Application of logic functions in mathematical modelation of neuron and 4. Aplication of binomial theorem and binomial distribution in genetics. Each application contains solved problems, a worksheet for students and a solution for each worksheet. Two application (1. and 2.) have been tested in teaching and as an assessment of my lessons students filled questionnaires. Results of these questionnaires are processed in the end of these chapters. This Thesis can be used in teaching or self-studying.

Prohlášení

Prohlašuji, že svoji diplomovou práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích, dne 2. prosince 2018 ....

Lucie Studená

Poděkování

Děkuji své školitelce RNDr. Ing. Janě Kalové, Ph.D. za odborné vedení diplomové práce, především za cenné rady a čas, který mi poskytla při konzultacích.

Obsah

1 Úvod **1**

2 Aplikace podmíněné pravděpodobnosti v lékařské diagnostice **3**

2.1 Motivace 3

2.2 Teorie podmíněné pravděpodobnosti 4

2.3 Lékařská diagnostika 6

2.4 Pracovní list 13

2.5 Realizovaná výuka 17

3 Aplikace exponenciální funkce v populační ekologii **25**

3.1 Motivace 26

3.2 Exponenciální a logaritmická funkce 28

3.3 Exponenciálně rostoucí populace 31

3.4 Pracovní list 41

3.5 Realizovaná výuka 47

4 Aplikace logických funkcí při matematickém modelování neuronu **55**

4.1 Motivace 56

4.2 Logické funkce a jejich grafické znázornění 58

4.3 Matematický model neuronu 63

4.4 Pracovní list 71

5 Aplikace binomické věty a binomického rozdělení v genetice **75**

5.1 Motivace 75

5.2 Binomická věta a binomické rozdělení 76

5.3 Pravděpodobnost narozeného potomstva určité skladby pohlaví v jedné rodině 82

5.4 Hardy-Weinbergova rovnováha ve velké panmiktické populaci 85

5.5 Pracovní list 90

6 Závěr **93**

Literatura **95**

A Článek ze sborníku konference Užití počítačů ve výuce matematiky 2017 **i**
A.1 Úvod ii
A.2 Podmíněná pravděpodobnost a Bayesův vzorec iii
A.3 Lékařská diagnostika v
A.3.1 Princip statistického vyhodnocování diagnostických testů v
A.3.2 Vliv prevalence na prediktivní hodnoty testu x
A.4 Návrh pracovního listu xii
A.5 Autorské řešení pracovního listu xv
A.6 Závěr xvi
Literatura xvii

.

1. Úvod

Matematika na středních školách bývá často vnímána jako náročný a ne příliš oblíbený předmět, což se mi potvrdilo nejen při mé pedagogické praxi. Studenti pokládají otázky typu: „K čemu mi matematika v životě vlastně bude, když chci studovat jazyky?“ Setkala jsem se nejednou také s názorem, že je v dnešní době plné techniky zbytečné se učit matematiku, neboť nám přece všechno dokáže spočítat počítač. Matematika však není izolovanou vědou, nýbrž nachází své široké uplatnění napříč mnohými vědními, lékařskými, technickými i ekonomickými obory, a její přínos v těchto vědách je často zcela klíčový. Troufám si tvrdit, že mnozí ze středoškolských studentů si tento fakt vůbec neuvědomují.

Mně známé učebnice matematiky bohužel nenabízí učitelům a studentům dostatek zajímavých aplikací matematiky, který by zdůrazňovaly její význam, a omezují se pouze na notoricky známé především fyzikální úlohy na např. poločas rozpadu, sílu působící na těleso apod.

Biologie se u studentů naopak těší poměrně velké oblibě možná právě proto, že si její přítomnost a využitelnost všude kolem nás dokáží studenti dobře představit (např. v lékařství). V biologii a jí příbuzných lékařských vědách lze najít několik zajímavých aplikací středoškolské matematiky. Proč tedy nepropojit tyto dvě vědní disciplíny a nepokusit se vzbudit ve studentech zájem pro matematiku pomocí aplikací z biologie a lékařství?

Ve své diplomové práci propojuji obory biologii a matematiku, které studuji jako dvou-aprobaci pro učitelství na střední škole. Cílem této práce je poskytnout učitelům středních škol návod, jak studentům ukázat využitelnost a nezbytnost matematiky, a zároveň jim přiblížit vybrané problematické kapitoly matematiky (podmíněnou pravděpodobnost, exponenciální a logaritmickou funkci, matematickou logiku, binomickou větu a binomické rozdělení). Dalším cílem práce je také nabídnout učitelům matematiky i biologie zajímavé náměty, jak obohatit výuku, zaujmout a motivovat studenty.

Každou vybranou aplikaci reprezentuje samostatná kapitola, která obsahuje stručné vysvětlení teoretického matematického i biologického základu, následuje představení principu samotné aplikace demonstrované na řešených a vysvětlených příkladech. Velkou výhodou aplikací je, že je lze využít jak v hodinách matematiky, tak i v hodinách biologie. Součástí práce je také příloha obsahující návrhy pracovních listů s autorským řešením, které slouží jako inspirace pro učitele a nabízí další úlohy k dané problematice.

Kapitoly jsou napsány tak, aby je mohli využít také učitelé matematiky, kteří nemají jako druhou aprobaci biologii, a naopak. Text lze ovšem také využít k samostudiu např. pro nadané studenty. Je pouze na uvážení každého učitele, v jakém rozsahu a obtížnosti aplikaci ve svých hodinách využije, a zda-li ji zařadí do běžné výuky, nebo ji představí jen vybrané skupině přírodovědně nadaných studentů, např. v rámci výběrového semináře.

Jednotlivé aplikace byly vybrány s ohledem na jejich náročnost tak, aby středoškolští studenti měli potřebný aparát matematiky i biologie a zároveň aby mohly sloužit jako inspirace učitelům pro rozšíření běžné výuky, popř. demonstraci vyučovaných teoretických poznatků.

V práci jsou použity jak příklady s reálnými převzatými údaji, tak i příklady s hypotetickými vlastními údaji. Zdroj reálných údajů je vždy v příkladu uveden.

2. Aplikace podmíněné pravděpodobnosti v lékařské diagnostice

Klíčová slova: Podmíněná pravděpodobnost, Bayesův vzorec, lékařská diagnostika, diagnostický test, specificita, senzitivita, prevalence.

Předpoklad: Znalost základů teorie pravděpodobnosti (viz např.: POLÁK, Josef. *Přehled středoškolské matematiky*. 10. vydání. Praha: Prometheus, 2015. ISBN 978-80-7196-458-2, s. 326 – 343).

V této kapitole se dozvíte:

- Co je to diagnostický test a kde se s ním můžeme setkat.
- Co je to podmíněná pravděpodobnost a jak se počítá. Jaký má tvar Bayesův vzorec a jak jej lze upravit pomocí věty o úplné pravděpodobnosti.
- Jací jsou ukazatelé správnosti diagnostického testu. Co je to specificita a senzitivita. Jak se provádí statistické vyhodnocování správnosti diagnostického testu.
- Co je to prevalence a jak její velikost souvisí s kvalitou diagnostického testu.

Použité značení:

A, B	jevy A, B
$A \cap B$	průnik jevů A, B
$A \cup B$	sjednocení jevů A, B
$P(A)$	pravděpodobnost jevu A
A'	opačný jev k jevu A

2.1 Motivace

Podmíněná pravděpodobnost a Bayesův vzorec nachází své využití mj. v populačních etiologických studiích[1] a v některých matematických modelech diagnostického, terapeutického či prognostického lékařského rozhodování. Používá se tedy v případech, kdy je potřeba vyhodnotit kvalitu screeningového nebo diagnostického testu.

[1] Etiologické studie jsou studie o příčinách a původech nemocí.

Tato kapitola se zabývá nejběžnější aplikací podmíněné pravděpodobnosti a Bayesova vzorce v lékařství – což je právě statistické vyhodnocování správnosti diagnostických testů. Diagnostické testy jsou speciální lékařské metody pomáhající s určitou pravděpodobností zjistit, zda pacient má, resp. nemá, zjišťovanou nemoc nebo určitou látku v těle.
Diagnostický test může představovat nejen určitou laboratorní metodu jako je histologické vyšetření, ale i např. ultrazvukové vyšetření nebo zátěžový kardiotest. Mezi diagnostické testy patří také testy volně zakoupitelné v lékárně – těhotenský test, test na alkohol nebo test potravinové intolerance. Na trhu existuje takovýchto různých testů nepřeberné množství (viz např. www.prozdravi.cz/diagnosticke-testy). Od diagnostických testů je vyžadována především jednoduchost, přijatelná cena, snadná a rychlá proveditelnost, ale také neškodnost a bezbolestnost.

Téma prezentované v této kapitole jsem zpracovala v článku s názvem Aplikace podmíněné pravděpodobnosti v lékařské diagnostice, který vyšel ve sborníku konference Užití počítačů ve výuce matematiky 2017. Článek je uvedený v celém rozsahu v příloze práce nebo je také dostupný na adrese http://home.pf.jcu.cz/~upvm/2017/wp-content/uploads/2018/02/Sbornik_UPVM_2017.pdf.

2.2 Teorie podmíněné pravděpodobnosti

Než si objasníme princip vyhodnocování kvality diagnostického testu, připomeneme si nezbytné znalosti z teorie podmíněné pravděpodobnosti.

S podmíněnou pravděpodobností souvisí pojem nezávislost jevů. Dva jevy jsou navzájem nezávislé, jestliže pravděpodobnost nastání jednoho z nich nezávisí na tom, zda-li druhý jev nastal nebo nenastal.

Věta 2.2.1. ([4], s. 15) [**Nezávislé jevy.**] Jsou-li A, B dva nezávislé jevy, pak pro pravděpodobnost jejich současného nastání platí

$$P(A \cap B) = P(A) \cdot P(B). \tag{2.1}$$

Číslo $P(A)$ označuje pravděpodobnost jevu A vzhledem k pevně danému souboru základních podmínek, po jejichž uskutečnění nastane jev A s pravděpodobností $P(A)$. Pokud k těmto základním podmínkám připojíme ještě další podmínku, že nastane jev B, tedy $P(B) > 0$, můžeme pak hledat pravděpodobnost jevu A za této podmínky. Takovéto pravděpodobnosti říkáme podmíněná pravděpodobnost.[3]

Definice 2.2.1. ([4], s. 14)[**Podmíněná pravděpodobnost.**] Podmíněnou pravděpodobnost jevu A vzhledem k jevu B označíme $P(A|B)$ a definujeme ji následovně:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \ P(B) > 0. \tag{2.2}$$

Uveďme si ilustrativní příklad. Mějme dva různé jevy A a B, jev A značí, že je určitý pacient určen jako pozitivní a jev B značí, že pacient skutečně trpí danou nemocí. $P(A)$ je pravděpodobnost, že je pacient určen jako pozitivní, $P(B)$ je pravděpodobnost, že pacient skutečně trpí nemocí. Podmíněná pravděpodobnost, že je *pacient určen jako pozitivní za předpokladu, že je skutečně nemocný* by se spočítala jako $P(A|B)$, kde $P(B) > 0$, dle vzorce (2.2) a pravděpodobnost, že je *pacient určen jako pozitivní za předpokladu, že nemocný není* by se dle stejného vzorce spočítala jako $P(A|B')$, kde $P(B') > 0$ a jev B' (pacient není nemocný) je jev opačný k jevu B, tj. $P(B') = 1 - P(B)$.

Jsou-li jevy A, B nezávislé, pak dosazením rovnosti (2.1) do vzorce pro výpočet podmíněné pravděpodobnosti (2.2) získáme:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

Platí tedy, že podmíněná pravděpodobnost vzhledem k nezávislému jevu se rovná nepodmíněné pravděpodobnosti. [4]

Ze vzorce pro podmíněnou pravděpodobnost (2.2) si vyjádříme $P(A \cap B)$ následujícím způsobem:

$$P(A \cap B) = P(A|B)P(B). \tag{2.3}$$

Pokud budeme chtít vyjádřit pravděpodobnost jevu B za podmínky nastoupení jevu A, zaměníme ve vzorci pro podmíněnou pravděpodobnost (2.2) jevy A a B

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \ P(A) > 0. \tag{2.4}$$

Rovnost (2.3) dosadíme do vzorce (2.4)

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}, \ P(A) > 0,$$

čímž jsme získali tzv. Bayesův vzorec pro dvojici jevů, který udává, jak podmíněná pravděpodobnost nějakého jevu souvisí s inverzní podmíněnou pravděpodobností.[5]

Věta 2.2.2. ([3], s. 31)**[Bayesův vzorec.]** Mějme dva náhodné jevy A a B s pravděpodobnostmi $P(A)$ a $P(B)$, $P(A) > 0$, potom platí

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}, \tag{2.5}$$

kde $P(A|B)$ je podmíněná pravděpodobnost jevu A za předpokladu, že nastal jev B.

Věta 2.2.3. ([3], s. 29)**[O úplné pravděpodobnosti.]** Nechť B, C jsou dva jevy, které se navzájem vylučují, přičemž jeden z nich nutně nastává. Potom pro libovolný jev A platí $A = (A \cap B) \cup (A \cap C)$ a pravděpodobnost $P(A)$ lze vyjádřit vztahem

$$P(A) = P(B)P(A|B) + P(C)P(A|C). \tag{2.6}$$

Bayesův vzorec (2.5) lze dále upravit dosazením vzorce (2.6) do jmenovatele za $P(A)$:[5]

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(B)P(A|B) + P(C)P(A|C)}. \tag{2.7}$$

2.3 Lékařská diagnostika

Princip statistického vyhodnocování diagnostických testů

Každý diagnostický test má dva možné výsledky, které nazýváme:

- **pozitivní výsledek**, ozn. A^+,
- **negativní výsledek**, ozn. A^-,

jedná se vlastně o dvojici opačných a tedy i navzájem neslučitelných jevů.
Dále je důležitá také skutečná přítomnost nemoci, tedy skutečnost, jestli je daná osoba opravdu nemocná, nebo není nemocná. Opět jsou to dva opačné navzájem neslučitelné jevy:

- **nemoc je přítomna**, ozn. H^+,
- **nemoc není přítomna**, ozn. H^-.[2]

Diagnostické schopnosti testu jsou při vyhodnocování prověřovány proti skutečnému stavu testovaných osob, kde se proti sobě srovnávají právě pozitivní (resp. negativní) výsledky testu a prokazatelně zjištěná přítomnost (resp. nepřítomnost) nemoci nebo látky. Pro takovéto

vyhodnocování je zavedeno několik ukazatelů správnosti (validity) představujících číselné ohodnocení testu ve vztahu k jeho chybovosti. Představíme si čtyři nejzákladnější z nich.[2]
První dva ukazatele správnosti diagnostických testů jsou tzv. *senzitivita testu* (schopnost testu rozpoznat nemocné osoby) a *specificita testu* (schopnost testu rozpoznat osoby bez nemoci). Tyto ukazatele definujeme pomocí podmíněné pravděpodobnosti:

- **senzitivita testu** je pravděpodobnost, že test bude pozitivní, když je osoba skutečně nemocná, tj. $P(A^+|H^+)$,
- **specificita testu** je pravděpodobnost, že test bude negativní, když je osoba zdravá, tj. $P(A^-|H^-)$.[2]

S pojmy senzitivita a specificita se lze běžně setkat v příbalovém letáku u volně zakoupitelných testů, nebo v popisu testu v internetové lékárně. Proto je dobré na tento fakt studenty upozornit nejlépe ukázkou např. příbalového letáku, aby si uvědomili praktický význam této aplikace.
Druhé dva ukazatele správnosti diagnostických testů jsou tzv. *prediktivní hodnoty*, které stejně jako předchozí ukazatele definujeme pomocí podmíněné pravděpodobnosti:

- **prediktivní hodnota pozitivního testu** je pravděpodobnost, že osoba skutečně je nemocná, když test vyjde pozitivní, tj. $P(H^+|A^+)$,
- **prediktivní hodnota negativního testu** je pravděpodobnost, že osoba skutečně není nemocná, když test vyjde negativní, tj. $P(H^-|A^-)$.[2]

Tabulka (2.1) objasňuje, jak lze ze zjištěných dat spočítat hodnoty zmíněných ukazatelů správnosti (jedná se vlastně o vypočítání podmíněné pravděpodobnosti na základě daných jevů – výsledků testu, a skutečné ověřené přítomnosti nemoci):

Tabulka 2.1: Obecná tabulka zjištěných hodnot diagnostického testu (upraveno podle [1]).

výsledek testu	**ověření přítomnosti nemoci / látky**		**celkem**
	přítomna (H^+)	***nepřítomna (H^-)***	
pozitivní (A^+)	a	b	a + b
negativní (A^-)	c	d	c + d
celkem	a + c	b + d	a + b + c + d

Z tabulky (2.1) je zřejmé, že a je počet lidí s pozitivním výsledkem testu, u nichž je nemoc přítomna, b je počet lidí s pozitivním výsledkem, u nichž je nemoc nepřítomna, c je počet lidí s negativním výsledkem, u nichž je nemoc přítomna a d je počet lidí s negativním výsledkem, u nichž je nemoc nepřítomna.

Hodnoty ukazatelů v procentech se poté určují pomocí podmíněné pravděpodobnosti jako:

$$P(A^+|H^+) = \frac{a}{a+c} \cdot 100, \quad P(A^-|H^-) = \frac{d}{b+d} \cdot 100,$$
$$P(H^+|A^+) = \frac{a}{a+b} \cdot 100, \quad P(H^-|A^-) = \frac{d}{c+d} \cdot 100. \text{ [1]}$$

Jako další ukazatel lze pro zajímavost zmínit ještě tzv. **celkovou správnost** S, která se definuje $S = \dfrac{a+d}{a+b+c+d}$. [6]

V následujícím jednoduchém příkladu si ukážeme určení ukazatelů správnosti diagnostických testů.

Příklad 2.3.1. Vypočítejte senzitivitu, specificitu a prediktivní hodnoty hypotetického testu na zjištění rakoviny na základně hodnot uvedených v tabulce (2.2).

Tabulka 2.2: Zjištěné hodnoty hypotetického testu na určení rakoviny.

výsledek testu	**ověření přítomnosti rakoviny**		**celkem**
	přítomna (H^+)	***nepřítomna (H^-)***	
pozitivní (A^+)	47	4	51
negativní (A^-)	5	38	43
celkem	52	42	94

Jev A^+ značí, že náhodně vybraný pacient má pozitivní test na rakovinu a opačný jev A^-, že osoba má negativní test na rakovinu. Jev H^+ značí, že je rakovina u pacienta skutečně přítomna a opačný jev H^-, že je rakovina nepřítomna. Určete, zda je test přesný (za přesný test zde považujte test s hodnotami ukazatelů vyššími než $80\,\%$).

Řešení:

Z tabulky (2.2) nejprve vypočítáme senzitivitu a specificitu testu:

$$P(A^+|H^+) = \frac{47}{52} \cdot 100 \doteq 90,4\,\%, \qquad P(A^-|H^-) = \frac{38}{42} \cdot 100 \doteq 90,5\,\%,$$

poté prediktivní hodnotu pozitivního a negativního testu:

$$P(H^+|A^+) = \frac{47}{51} \cdot 100 \doteq 92,2\,\%, \qquad P(H^-|A^-) = \frac{38}{43} \cdot 100 \doteq 88,4\,\%.$$

Vidíme, že daný test je poměrně kvalitní a přesný, neboť se hodnoty ukazatelů pohybují nad $80\,\%$. Pravděpodobnost, že je pacient skutečně nemocný, když mu test vyjde pozitivní, je dokonce $92,2\,\%$.

Senzitivita a specificita jsou především populační ukazatele, protože se zjišťují na základě znalosti skutečné přítomnosti (resp. nepřítomnosti) onemocnění u určitého souboru lidí. Zajímají zejména lékaře, kteří je pak mohou použít pro porovnání diagnostické správnosti dvou různých testů. U konkrétního pacienta, který přijde do ordinace s výsledky testu, skutečnou přítomnost (resp. nepřítomnost) nemoci lékař nezná. Pro pacienty jsou zajímavější ukazatele prediktivní hodnoty, které vycházejí z určitého výsledku jejich testu, a které v případě, že jejich vlastní test vyšel pozitivně (resp. negativně), určují s jakou pravděpodobností danou nemoc skutečně mají (resp. nemají). [2]
Stejně tak jako v jiných statistických metodách, i u těchto testů zvyšuje jejich kvalitu dostatečná velikost experimentálního vzorku (tj. velikost souboru testovaných lidí). Pokud by byl tento soubor malý, poroste pravděpodobnost, že někteří specifičtí pacienti nebudou zachyceni a odhady specificity a senzitivity budou zkreslené. Mezní hodnoty senzitivity a specificity určující, zda je daný test kvalitní, resp. kvalitnější než jiné testy, které jsou aktuálně dostupné, se určují obtížně. Závisí totiž na míře poznání dané oblasti a na dosažitelné správnosti dostupných testů – někdy jsou pro nás hodnoty nad 60 % „výhrou“, zatímco v jiné oblasti se tyto hodnoty blíží ke 100 %.[2]

V dalším příkladu budeme porovnávat dva těhotenské testy. Těhotenské testy fungují na principu zjištění přítomnosti hormonu gonadotropinu v moči nebo v krvi. Gonadotropin začíná být produkován zárodkem placenty v momentě oplodnění. S vývinem placenty jeho koncentrace v moči či krvi v prvních týdnech po otěhotnění výrazně stoupá téměř každým dnem.

Příklad 2.3.2. Dva hypotetické těhotenské testy – Mimicheck a Pregnus – na svém obalu tvrdí, že jsou nejlepším volně dostupným těhotenským testem. Úkolem je určit, který z nich je lepší (tj. spočítat a porovnat ukazatele správnosti).
K dispozici jsou výsledky obou těchto testů a výsledky ověření těhotenství, které byly zjišťovány na stejném souboru 200 žen (z toho 66 žen bylo lékařským vyšetřením ověřeno jako skutečně těhotných), zanesených v tabulce (2.3):

Tabulka 2.3: Zjištěné hodnoty pro Mimicheck a Pregnus.

výsledek testu Mimicheck	**ověření skutečného těhotenství**		**celkem**
	těhotná (H^+)	***není těhotná (H^-)***	
pozitivní (A^+)	63	4	67
negativní (A^-)	3	130	133
celkem	66	134	200

výsledek testu Pregnus	**ověření skutečného těhotenství**		**celkem**
	těhotná (H^+)	***není těhotná (H^-)***	
pozitivní (A^+)	60	2	62
negativní (A^-)	6	132	138
celkem	66	134	200

Řešení:

Nejdříve vyhodnotíme Mimicheck:

- senzitivita ... $P(A^+|H^+) = \frac{63}{66} \cdot 100 \doteq 95,5\ \%$
- specificita ... $P(A^-|H^-) = \frac{130}{134} \cdot 100 \doteq 97,0\ \%$
- prediktivní hodnota pozitivního testu ... $P(H^+|A^+) = \frac{63}{67} \cdot 100 \doteq 94,0\ \%$
- prediktivní hodnota negativního testu ... $P(H^-|A^-) = \frac{130}{133} \cdot 100 \doteq 97,7\ \%$

Poté vyhodnotíme Pregnus:

- senzitivita ... $P(A^+|H^+) = \frac{60}{66} \cdot 100 \doteq 90,9\ \%$
- specificita ... $P(A^-|H^-) = \frac{132}{134} \cdot 100 \doteq 98,5\ \%$
- prediktivní hodnota pozitivního testu ... $P(H^+|A^+) = \frac{60}{62} \cdot 100 \doteq 96,8\ \%$
- prediktivní hodnota negativního testu ... $P(H^-|A^-) = \frac{132}{138} \cdot 100 \doteq 95,7\ \%$

Porovnáním obou testů zjišťujeme, že oba testy vykazují vysoké pravděpodobnosti nad $90\ \%$ a jsou tedy oba poměrně kvalitní.[2] Mimicheck má vyšší senzitivitu, což znamená, že s větší pravděpodobností rozpozná skutečně těhotné ženy. Pregnus má o něco větší specificitu a je tedy v porovnání s Mimicheckem citlivější na vyloučení žen, které nejsou těhotné. Navíc Pregnus vykazuje větší pravděpodobnost $(96,8\ \%)$, že je žena skutečně těhotná, když test vyjde pozitivní, což je většinou hledisko, které nejvíce zajímá právě ženy provádějící těhotenské testy.

[2]Ve skutečnosti příbalové letáky těhotenských testů uvádějí až $100\ \%$ hodnoty senzitivity a specificity. V zadání tohoto příkladu byly hodnoty záměrně sníženy, aby bylo možné testy porovnat.

Vliv prevalence na prediktivní hodnoty testu

Prediktivní hodnoty pozitivního a negativního testu mj. úzce souvisí s **prevalencí** určité sledované nemoci (nebo obecně nějakého defektu) v dané populaci. V určitém časovém okamžiku lze prevalenci vyjádřit jako procento pacientů se sledovanou nemocí určené ze všech osob v dané populaci.[2]

Jak moc mohou prediktivní hodnoty záviset na velikosti prevalence si ukážeme na řešeném příkladu se záměrně vybranými dvěma populacemi s extrémně rozdílnými počty nemocných jedinců.

Než se začneme zabývat příkladem, musíme si vyjádřit prediktivní hodnoty pomocí prevalence. Označme prevalenci jako $P(H^+)$, tj. pravděpodobnost, že určitá osoba je skutečně nemocná (v daný okamžik a v dané populaci) a platí $1 - P(H^+) = P(H^-)$, kde $P(H^-)$ je pravděpodobnost, že určitá osoba není nemocná.

Prediktivní hodnoty si nyní vyjádříme pomocí upraveného Bayesova vzorce (2.7) s úplnou pravděpodobností a pomocí hodnot senzitivity a specificity:[2]

$$P(H^+|A^+) = \frac{P(A^+|H^+)P(H^+)}{P(A^+|H^+)P(H^+) + P(A^+|H^-)P(H^-)},$$

$$P(H^-|A^-) = \frac{P(A^-|H^-)P(H^-)}{P(A^-|H^-)P(H^-) + P(A^-|H^+)P(H^+)}.$$

Příklad 2.3.3. Úkolem je určit a porovnat pozitivní a negativní prediktivní hodnoty hypotetického testu na zjištění HIV pozitivity, u kterého je výrobcem garantována senzitivita $98\ \%$ a specificita $99\ \%$ pro populaci Lesotha, kde se k 1. 1. 2017 vyskytuje $25\ \%$ HIV pozitivních a poté pro populaci České republiky, kde se k 1. 1. 2017 vyskytuje pouze $0,03\ \%$ HIV pozitivních.[3]

Řešení:

Nejprve budeme uvažovat populaci Lesotha s relativně vysokou prevalencí HIV pozitivity:

- $P(H^+) = 0,25$:

$$P(H^+|A^+) = \frac{0,98 \cdot 0,25}{0,98 \cdot 0,25 + (1 - 0,99) \cdot (1 - 0,25)} \doteq 0,970$$

$$P(H^-|A^-) = \frac{0,99 \cdot (1 - 0,25)}{0,99 \cdot (1 - 0,25) + (1 - 0,98) \cdot 0,25} \doteq 0,993$$

[3] Jedná se o zaokrouhlené hodnoty převzaté z reálných statistik – pro ČR ze Státního zdravotního ústavu: http://www.szu.cz/tema/prevence/zprava-o-vyskytu-a-sireni-hiv-aids-za-rok-2017, pro Lesotho z http://www.unaids.org/en/regionscountries/countries/lesotho.

A poté populaci České republiky s relativně nízkou prevalencí HIV pozitivity:

- $P(H^+) = 0,0003$:

$$P(H^+|A^+) = \frac{0,98 \cdot 0,0003}{0,98 \cdot 0,0003 + (1-0,99) \cdot (1-0,0003)} \doteq 0,029$$

$$P(H^-|A^-) = \frac{0,99 \cdot (1-0,0003)}{0,99 \cdot (1-0,0003) + (1-0,98) \cdot 0,0003} \doteq 0,999.$$

Je patrné, že v populaci s relativně vysokou prevalencí HIV pozitivity má kvalitní test vysokou pozitivní i negativní prediktivní hodnotu, tj. osoby s pozitivním (resp. negativním) testem mají vysokou pravděpodobnost, že jsou skutečně HIV pozitivní (resp. negativní).

Ovšem v populaci s nízkou prevalencí HIV pozitivity, má kvalitní test sice vysokou negativní prediktivní hodnotu – na $99,9$ % jsou osoby s negativním výsledkem testu opravdu HIV negativní, ale relativně nízkou pozitivní prediktivní hodnotu – pouze na $2,9$ % jsou osoby s pozitivním testem skutečně HIV pozitivní.

Kvalita testu daná vysokou senzitivitou a specificitou je tedy velice relativní a hodně záleží na velikosti prevalence uvažované nemoci v této populaci.

2.4 Pracovní list

Téma: Aplikace podmíněné pravděpodobnosti v lékařské diagnostice

Jméno a třída: Datum:

Příklad č. 1 Hypotetický drogový test D7 slouží pro kvalitativní stanovení drogových metabolitů v lidské moči. Jedna z testem detekovaných látek je THC (Tetrahydrocannabinol, což je hlavní psychoaktivní látka nacházející se v konopí, mj. pochází z drog marihuana a hašiš). Detekční mez uvedená výrobcem je pro THC-kannabinol, 11-nor-8-THC-9-COOH a 11-nor-9-THC-COOH je 50 ng/ml. Výrobce udává $92,8\ \%$ specificitu a $95\ \%$ senzitivitu.

Vysvětli pomocí pravděpodobnosti, co znamená specificita a senzitivita na tomto konkrétním diagnostickém testu:

Na základě udaných hodnot senzitivity a specificity rozhodni, zda je možné test pokládat za kvalitní a přesný:

Příklad č. 2 Hypotetický samodiagnostický test potravinové intolerance AlergSeeker je velmi jednoduchý a rychlý. Testovací podložka s nanesenými potravinovými extrakty během 40 minut ukáže, na kterou z potravin má člověk vytvořené protilátky (tj. je alergický). Stačí k tomu pouze malá kapička krve ze špičky prstu, která se rozpustí v roztoku. Tak lze identifikovat přítomnost protilátek na jednu či více potravin v závislosti na výskytu modře zbarvených bodů na reakční podložce. Mezi testované potraviny patří obiloviny, luštěniny, ořechy, maso a ryby, zelenina a ovoce, vejce, kravské mléko aj.

V následující tabulce jsou uvedeny zjištěné hodnoty pro testování přítomnosti protilátek na laktózu v kravském mléce a na gluten (lepek) v obilovinách. Obě testování byla provedena na souboru 300 lidí:

výsledek testu (laktóza)	**ověření skutečné intolerance**		**celkem**
	přít. protilátek (H^+)	***nepř. protilátek (H^-)***	
pozitivní (A^+)	44	1	45
negativní (A^-)	1	254	255
celkem	45	255	300

výsledek testu (lepek)	**ověření skutečné intolerance**		**celkem**
	přít. protilátek (H^+)	***nepř. protilátek (H^-)***	
pozitivní (A^+)	19	0	19
negativní (A^-)	2	279	281
celkem	21	279	300

Vypočítej všechny ukazatele správnosti u obou testování a porovnej prediktivní hodnotu pozitivního testu (pravděpodobnost, že osoba je skutečně alergická, když test vyjde pozitivní) při zjišťování přítomnosti protilátek na kravské mléko a na lepek.

Příklad č. 3 V následující tabulce jsou uvedeny čtyři neinvazivní zátěžové kardiotesty, což jsou testy činnosti srdce. Neinvazivní znamená, že se při jejich provádění nezasahuje dovnitř organismu. Doplň chybějící data v tabulce a na základě senzitivity a specificity porovnej kvalitu těchto testů.

	druh testu	počet pacientů	zjištěné hodnoty			senzitivita (%)	specificita (%)
			výsledek testu	ověření přítomnosti nemoci: *přítomna* (H^+)	*nepřítomna* (H^-)		
1.	**zátěžová EKG**	**11671**	*pozitivní* (A^+)		2345		**71,2**
			negativní (A^-)	1047	5790		
2.	**zátěžový SPECT**	**3343**	*pozitivní* (A^+)	1024	641		
			negativní (A^-)	433	1245		
3.	**zátěžová echokardiografie**		*pozitivní* (A^+)	360			**74,2**
			negativní (A^-)	59	224		
4.	**zátěžová magnetická rezonance**		*pozitivní* (A^+)	294		**90,2**	**85,2**
			negativní (A^-)		271		

Řešení pracovního listu:

Příklad č. 1 Senzitivita drogového testu $95\ \%$ znamená pravděpodobnost $0,95$, že tento test bude pozitivní, když má osoba provádějící test skutečně v těle takové množství THC, které je test schopný detekovat. Specificita drogového testu $92,8\ \%$ znamená pravděpodobnost $0,928$, že tento test bude negativní, když nemá osoba provádějící test v těle žádné množství THC nebo když má v těle nižší množství THC, než je test schopný detekovat.

Na základě hodnot senzitivity a specificity, které jsou vyšší než $90\ \%$, lze usoudit, že je test relativně kvalitní a přesný.

Příklad č. 2 Intolerance laktózy: $P(A^+|H^+) \doteq 97,8\ \%$, $P(A^-|H^-) \doteq 99,6\ \%$, $P(H^+|A^+) \doteq 97,8\ \%$, $P(H^-|A^-) \doteq 99,6\ \%$.

Intolerance lepku: $P(A^+|H^+) \doteq 90,5\ \%$, $P(A^-|H^-) = 100\ \%$, $P(H^+|A^+) = 100\ \%$, $P(H^-|A^-) \doteq 99,3\ \%$.

Test vykazuje u zjišťování intolerance laktózy i lepku vysoké hodnoty. Prediktivní hodnota pozitivního testu je vyšší u zjišťování alergie na lepek (100 %).

Příklad č. 3 Hodnoty lze snadno dopočítat z definic senzitivity a specificity.

<table>
<tr><th colspan="2">druh testu</th><th>počet pacientů</th><th colspan="3">zjištěné hodnoty</th><th>senzitivita (%)</th><th>specificita (%)</th></tr>
<tr><td rowspan="4">1.</td><td rowspan="4">zátěžová EKG</td><td rowspan="4">11671</td><td rowspan="2">výsledek testu</td><td colspan="2">ověření přítomnosti nemoci</td><td rowspan="4">70,4</td><td rowspan="4">71,2</td></tr>
<tr><td>přítomna (H^+)</td><td>nepřítomna (H^-)</td></tr>
<tr><td>pozitivní (A^+)</td><td>2489</td><td>2345</td></tr>
<tr><td>negativní (A^-)</td><td>1047</td><td>5790</td></tr>
<tr><td rowspan="4">2.</td><td rowspan="4">zátěžový SPECT</td><td rowspan="4">3343</td><td rowspan="2">výsledek testu</td><td colspan="2">ověření přítomnosti nemoci</td><td rowspan="4">70,3</td><td rowspan="4">66</td></tr>
<tr><td>přítomna (H^+)</td><td>nepřítomna (H^-)</td></tr>
<tr><td>pozitivní (A^+)</td><td>1024</td><td>641</td></tr>
<tr><td>negativní (A^-)</td><td>433</td><td>1245</td></tr>
<tr><td rowspan="4">3.</td><td rowspan="4">zátěžová echokardiografie</td><td rowspan="4">721</td><td rowspan="2">výsledek testu</td><td colspan="2">ověření přítomnosti nemoci</td><td rowspan="4">85,9</td><td rowspan="4">74,2</td></tr>
<tr><td>přítomna (H^+)</td><td>nepřítomna (H^-)</td></tr>
<tr><td>pozitivní (A^+)</td><td>360</td><td>78</td></tr>
<tr><td>negativní (A^-)</td><td>59</td><td>224</td></tr>
<tr><td rowspan="4">4.</td><td rowspan="4">zátěžová magnetická rezonance</td><td rowspan="4">644</td><td rowspan="2">výsledek testu</td><td colspan="2">ověření přítomnosti nemoci</td><td rowspan="4">90,2</td><td rowspan="4">85,2</td></tr>
<tr><td>přítomna (H^+)</td><td>nepřítomna (H^-)</td></tr>
<tr><td>pozitivní (A^+)</td><td>294</td><td>47</td></tr>
<tr><td>negativní (A^-)</td><td>32</td><td>271</td></tr>
</table>

Jako nejméně kvalitní zátěžový kardiotest se jeví SPECT, naopak jako nejvíce kvalitní se jeví magnetická rezonance.

2.5 Realizovaná výuka

Dne 23. 2. 2018 jsem odučila na Gymnáziu Jírovcova v Českých Budějovicích dvě vyučovací hodiny na téma Aplikace podmíněné pravděpodobnosti v lékařské diagnostice. Výuka proběhla v 7. ročníku osmiletého cyklu a byla pojata jako opakování podmíněné pravděpodobnosti a obohacení učiva o zajímavou aplikaci z lékařství. Jedná se o třídu s matematicky nadanými studenty, neboť průměr třídy ze známek z matematiky na posledním vysvědčení byl $1,76$. Ve třídě bylo v době mé výuky přítomno 29 studentů.

Studenti byli požádáni o vyplnění krátkého anonymního dotazníku, který zjišťoval vztah studentů k matematice a souvislost známek a obliby matematiky, k příkladům na aplikace matematiky v jiných oborech, jejich názor na odučenou dvouhodinovku a také na povinnou maturitu z matematiky. Dotazník se skládal ze dvou částí – první část studenti vyplnili před začátkem výuky, druhou po jejím skončení. Dotazník vypadal následovně:

1. část

1. **Baví tě matematika, popř. některá její část?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

2. **Jakou jsi měl(a) známku z matematiky na posledním vysvědčení?**

   1 – 2 – 3 – 4 – 5

3. **Počítáte v matematice příklady na využití matematiky ve fyzice, technice, ... ?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

4. **Uvítal(a) bys více příkladů na využití matematiky v přírodních vědách nebo v technice?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

2. část

1. **Přišla ti vyučovací hodina na téma Aplikace podmíněné pravděpodobnosti v lékařské diagnostice zajímavá?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

2. **Pomohl ti výklad a příklady v této hodině pochopit podmíněnou pravděpodobnost?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

3. **Uvítal(a) bys více takových hodin, které by se věnovaly využití matematiky v přírodních vědách?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

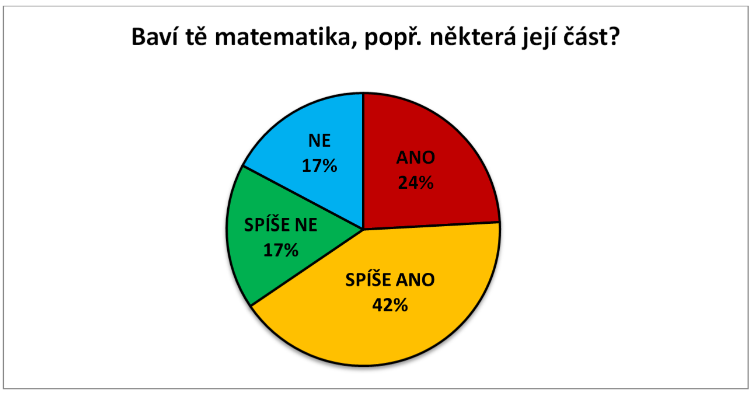
4. **Jaký je tvůj názor na povinnou maturitu z matematiky?** (otevřená odpověď)

Osnova výuky byla následující:

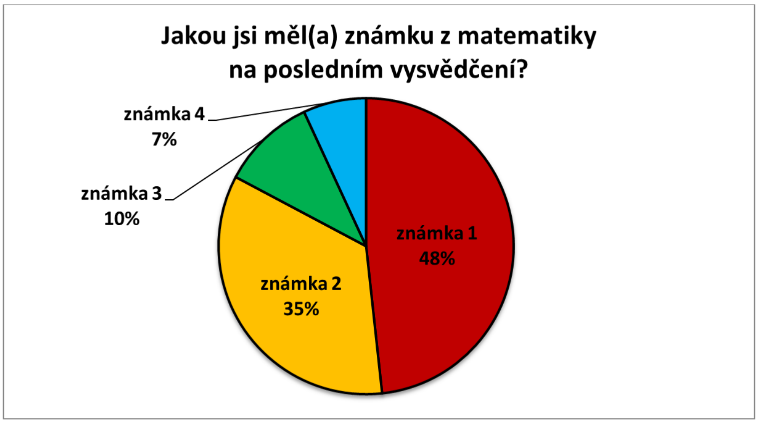
- vyplnění 1. části dotazníku, rozdání pracovního listu (viz podkapitola 2.4), představení cílů výuky,
- vysvětlení pojmů lékařská diagnóza, diagnostika, test; představení vybraných diagnostických testů a jejich principu,
- zopakování vzorce pro podmíněnou pravděpodobnost,
- vysvětlení principu lékařské diagnostiky – jevy v lékařské diagnostice, ukazatele správnosti, společné vyřešení vzorového příkladu na tabuli (viz Příklad 2.3.1),
- zopakování vzorce pro úplnou pravděpodobnost, Bayesova vzorce a jeho odvození, Bayesova vzorce s úplnou pravděpodobností,
- vysvětlení pojmu prevalence, společné vyřešení vzorového příkladu demonstrujícího vliv prevalence na prediktivní hodnoty testu (viz Příklad 2.3.3),
- společné vypracování 1. příkladu z pracovního listu a zadání zbylých dvou příkladů za dobrovolný domácí úkol,
- vyplnění 2. části dotazníku.

Zpracovaná data z dotazníku

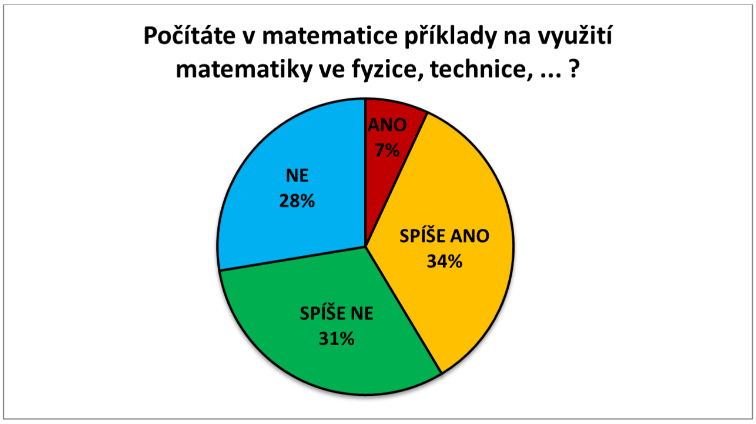
U první otázky mne velmi mile překvapilo, že 66 % studentů uvedlo, že je matematika, popř. některá její část, baví nebo spíše baví. Což je poněkud opačná situace, než se kterou jsem se doposud setkala.



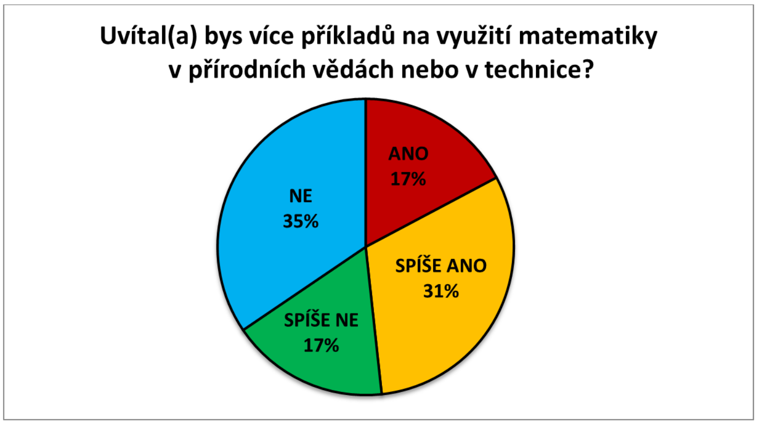
Fakt, že je v této třídě mnoho studentů se zálibou v matematice, podporuje druhý graf, který ukazuje procentuální skladbu známek z matematiky na posledním vysvědčení. Celkový průměr třídy z těchto známek byl $1,76$. Téměř polovina studentů dostala známku 1 a více jak třetina studentů známku 2. Lze tedy studenty této třídy právem považovat za matematicky nadané. Není asi žádným překvapením, že ti studenti, kteří z matematiky na vysvědčení měli 3 nebo 4 uvedli, že je matematika nebaví nebo spíše nebaví. Zajímavý je však poznatek, že dva studenti, kteří uvedli, že je matematika nebaví nebo spíše nebaví, z ní měli na vysvědčení 1.



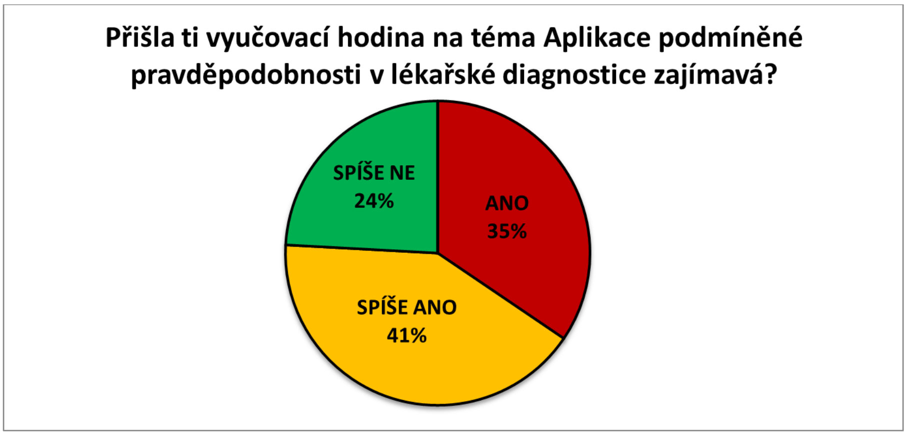
Další dvě otázky zjišťovaly, jestli se studenti domnívají, že počítají dostatek příkladů na využití matematiky v přírodních vědách a technice, a jestli by uvítali více takových příkladů. Ze všech dotazovaných se $41\ \%$ domnívá, že počítají dostatek takových příkladů a $59\ \%$ si naopak myslí, že těchto příkladů mnoho nepočítají.



Dle uvedených odpovědí by $48\ \%$ studentů ocenilo více příkladů na využití matematiky v přírodních vědách a technice, ovšem $52\ \%$ o více takových příkladů nestojí.

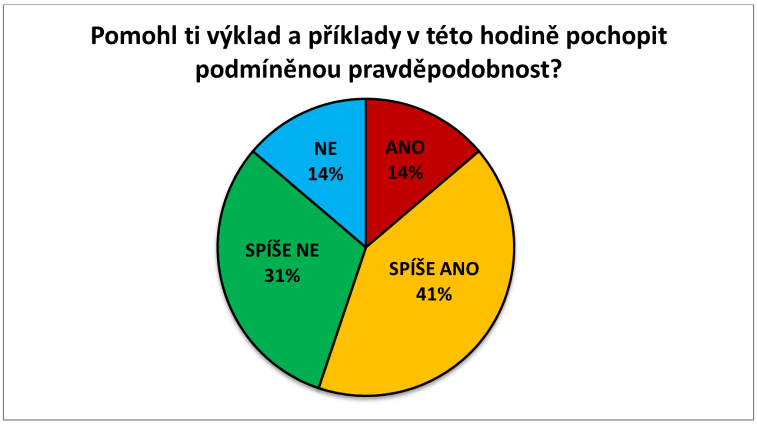


Příjemným zjištěním pro mne bylo, že 76 % studentů zhodnotilo moji „dvouhodinovku“ jako zajímavou, či spíše zajímavou, a žádný ze studentů neuvedl, že by ho výklad vyloženě nezaujal. Doufám, že studenti odpovídali zcela upřímně a že neodpovídali kladně jen proto, aby mne potěšili.

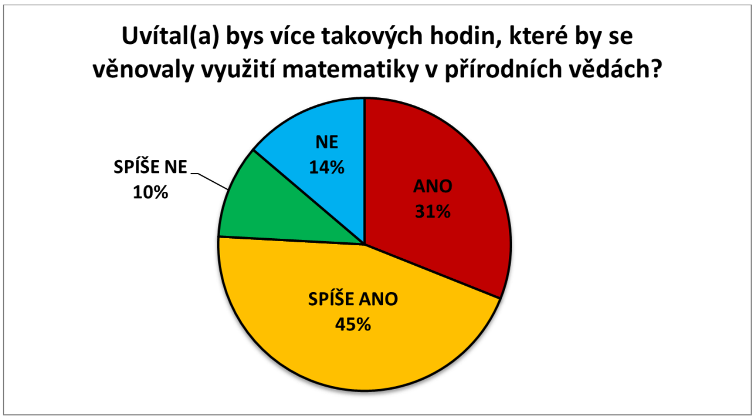


Zájem studentů o problematiku diagnostických testů byl vidět především, když jsem jim ukazovala fotografie některých zajímavých diagnostických testů, popisovala k čemu slouží, jaký je princip jejich provádění a fungování. Také je velmi zaujala ukázka příbalového letáku jednoho domácího diagnostického testu, který demonstroval fakt, že se s uvedenými pojmy z lékařské diagnostiky mohou setkat v běžném životě.

Výuka byla zaměřená na opakování podmíněné pravděpodobnosti, kterou měli studenti již probranou a tedy by jí měli již rozumět, proto mne nijak nepřekvapilo, že téměř polovina dotazovaných uvedla, že jim představení aplikace podmíněné pravděpodobnosti v lékařské diagnostice nepomohlo problematiku podmíněné pravděpodobnosti pochopit. Druhá polovina ovšem naopak uvedla, že jim můj výklad k jejímu pochopení pomohl. Potěšující je fakt, že tak odpověděli právě 3 studenti se známkou z matematiky horší než 2.



Velmi mne potěšilo také zjištění, že 76 % dotazovaných by uvítalo nebo spíše uvítalo více takovýchto hodin. Jsem ráda, že moje snaha o propojení matematiky a biologie byla úspěšná a že se studentům můj výklad líbil.



V poslední (otevřené) otázce jsem zjišťovala názor studentů na povinnou maturitu z matematiky. 62 % bylo proti a 38 % bylo pro povinnou maturitu. Zde většinou ti studenti, kteří byli proti povinné maturitě z matematiky, argumentovali, že ne každý se chce po maturitě věnovat matematice a že ve spoustě zaměstnání není matematika potřeba, a proto se jim zdá zbytečné, aby museli všichni skládat povinně maturitní zkoušku z matematiky. Uvedli, že je důležité, aby se každý zaměřil na předměty, které jej zajímají a je v nich dobrý. Podle nich je správné, aby byla maturita z matematiky volitelná. Některým studentům povinná maturita nevadí, ale dodávali, že by měla mít nastavenou vyšší úroveň. Někteří studenti, kteří byli pro povinnou maturitu, napsali, že je matematika užitečná v životě každého, neboť má široké uplatnění, a proto by z ní měla být maturita povinná.

Následující tabulka ukazuje četnosti jednotlivých odpovědí na dané otázky z obou částí dotazníku:

Tabulka 2.4: Četnosti jednotlivých odpovědí na otázky z obou částí dotazníku.

Četnosti jednotlivých odpovědí z dotazníku								
odpovědi	část 1				část 2			
	otázka 1	otázka 2	otázka 3	otázka 4	otázka 1	otázka 2	otázka 3	otázka 4
ANO	7		2	5	10	4	9	
SPÍŠE ANO	12		10	9	12	12	13	
SPÍŠE NE	5		9	5	7	9	3	
NE	5		8	10	0	4	4	
známka 1		14						
známka 2		10						
známka 3		3						
známka 4		2						
známka 5		0						
PRO								11
PROTI								18
celkem	29	29	29	29	29	29	29	29

V poslední tabulce je uveden přehled všech dotazníků (označených číslem) a odpovědí na jednotlivé otázky. Lze si tedy například všimnout, jak se změnil názor studentů na příklady demonstrující aplikace matematiky v přírodních vědách, technice apod. Celkem 8 dotazovaných v 4. otázce 1. části dotazníku uvedli, že by neuvítali (resp. spíše neuvítali) více takových příkladů, avšak po odučené dvouhodinovce ve 3. otázce 2. části dotazníku odpověděli, že by uvítali (resp. spíše uvítali) více takovýchto hodin, které by se věnovali aplikacím matematiky v přírodních vědách.

Tabulka 2.5: Přehled odpovědí u jednotlivých dotazníků.

číslo dotazníku	část 1				část 2			
	otázka 1	otázka 2	otázka 3	otázka 4	otázka 1	otázka 2	otázka 3	otázka 4
1	ANO	2	SPÍŠE ANO	ANO	ANO	ANO	ANO	PROTI
2	SPÍŠE ANO	1	ANO	ANO	SPÍŠE NE	SPÍŠE NE	SPÍŠE ANO	PRO
3	NE	2	SPÍŠE ANO	ANO	SPÍŠE ANO	SPÍŠE NE	ANO	PRO
4	SPÍŠE ANO	1	NE	NE	ANO	SPÍŠE NE	SPÍŠE NE	PROTI
5	SPÍŠE ANO	1	NE	NE	SPÍŠE NE	NE	ANO	PROTI
6	ANO	2	SPÍŠE ANO	SPÍŠE NE	ANO	ANO	ANO	PROTI
7	SPÍŠE ANO	1	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	ANO	ANO	PROTI
8	ANO	1	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	ANO	PROTI
9	SPÍŠE ANO	1	SPÍŠE NE	SPÍŠE NE	ANO	ANO	SPÍŠE ANO	PROTI
10	NE	2	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	PROTI
11	SPÍŠE ANO	2	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	PRO
12	SPÍŠE ANO	1	SPÍŠE ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	ANO	PROTI
13	SPÍŠE NE	2	SPÍŠE ANO	SPÍŠE NE	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	PRO
14	SPÍŠE ANO	1	SPÍŠE NE	SPÍŠE ANO	ANO	SPÍŠE NE	SPÍŠE ANO	PROTI
15	SPÍŠE NE	1	NE	NE	SPÍŠE NE	SPÍŠE NE	SPÍŠE NE	PROTI
16	NE	1	NE	NE	SPÍŠE NE	NE	NE	PROTI
17	SPÍŠE NE	3	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	PROTI
18	ANO	1	SPÍŠE ANO	SPÍŠE ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	PRO
19	SPÍŠE NE	3	ANO	NE	ANO	SPÍŠE ANO	NE	PROTI
20	SPÍŠE ANO	2	NE	NE	SPÍŠE ANO	SPÍŠE ANO	NE	PROTI
21	NE	4	SPÍŠE ANO	NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	PROTI
22	SPÍŠE ANO	2	NE	NE	SPÍŠE ANO	SPÍŠE ANO	NE	PRO
23	ANO	1	SPÍŠE ANO	ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	PRO
24	ANO	1	SPÍŠE NE	NE	SPÍŠE NE	SPÍŠE NE	SPÍŠE ANO	PROTI
25	SPÍŠE ANO	2	NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	PRO
26	NE	3	NE	SPÍŠE ANO	ANO	SPÍŠE ANO	ANO	PRO
27	SPÍŠE ANO	2	SPÍŠE NE	SPÍŠE NE	SPÍŠE NE	NE	SPÍŠE NE	PRO
28	ANO	1	SPÍŠE NE	SPÍŠE ANO	ANO	SPÍŠE NE	SPÍŠE ANO	PRO
29	SPÍŠE NE	4	SPÍŠE ANO	NE	SPÍŠE ANO	NE	ANO	PROTI

Na závěr bych chtěla dodat, že jsem byla velmi spokojená s reakcemi studentů na můj výklad a že doufám v to, že představení diagnostických testů a principu ověřování jejich správnosti jim bude v životě užitečné.

3. Aplikace exponenciální funkce v populační ekologii

Klíčová slova: Populace, populační dynamika, exponenciální a logaritmická funkce, exponenciální růst, exponenciálně rostoucí populace.

Předpoklad: Základní znalost funkcí, jejich vlastnosti, funkce inverzní, exponenciální, logaritmická (viz např.: POLÁK, Josef. *Přehled středoškolské matematiky*. 10. vydání. Praha: Prometheus, 2015. ISBN 978-80-7196-458-2, od s. 130).

V této kapitole se dozvíte:

- Co je to populační dynamika, čím jsou populace limitovány, za jakých podmínek roste populace exponenciálně a jaké jsou příklady exponenciálně rostoucích populací.
- Jak je definována exponenciální a logaritmická funkce, jaké jsou jejich základní vlastnosti a grafy.
- Jak lze exponenciální růst matematicky modelovat pro různé typy populací a jak můžeme určit velikost populace za určitý čas.
- Kde mimo populační ekologii lze v biologii uplatnit exponenciální růst.

Použité značení:

K	nosná kapacita prostředí
$f(x),\ f^{-1}(x)$	funkce f, inverzní funkce k funkci f
$D_f,\ H_f$	definiční obor funkce f, obor hodnot funkce f
N	velikost populace
α	konečná růstová rychlost
n	koeficient natality
m	koeficient mortality
r	relativní velikost změny
$f'(x),\ \dfrac{\mathrm{d}f(x)}{\mathrm{d}x}$	derivace funkce f podle proměnné x

3.1 Motivace

Populaci chápejme jako soubor jedinců stejného druhu vyskytující se na určitém místě v konkrétním okamžiku.[8]

Jedná se tedy například o populaci obyvatel České republiky v roce 2017 nebo také o populaci bakterií *Escherichia coli* ve střevní mikroflóře pana X. ke dni 27. 6. 2018.

Populacemi, změnami jejich hustoty, strukturou apod. se zabývá **populační ekologie** (demekologie). Jakékoli změny počtu jedinců v populaci se označují jako **dynamika populace**. Nejvíce je populační dynamika ovlivňována **natalitou** (produkcí mláďat), **mortalitou** (úmrtností jedinců) a **migracemi** – imigrací a emigrací (tj. přesuny jedinců do a z populace).[8]

Vše, co je organismy spotřebováváno za účelem růstu, rozmnožování a zachování života, se označuje jako **zdroje** (např. potrava, životní prostor, voda, živiny). O zdroje, kterých není nikdy dostatek pro všechny, organismy mezi sebou soupeří (kompetitují). **Kompetice o zdroje** může mít na populaci negativní limitující vliv. Může dojít ke snížení hustoty populace, či dokonce k jejímu vymírání.[9]

Například v důsledku nedostatku stromových hnízdních dutin v určitém lese může dojít k nižší reprodukci a následně ke snížení počtu jedinců populace sýkor modřinek (*Cyanistes caeruleus*) tohoto lesa.

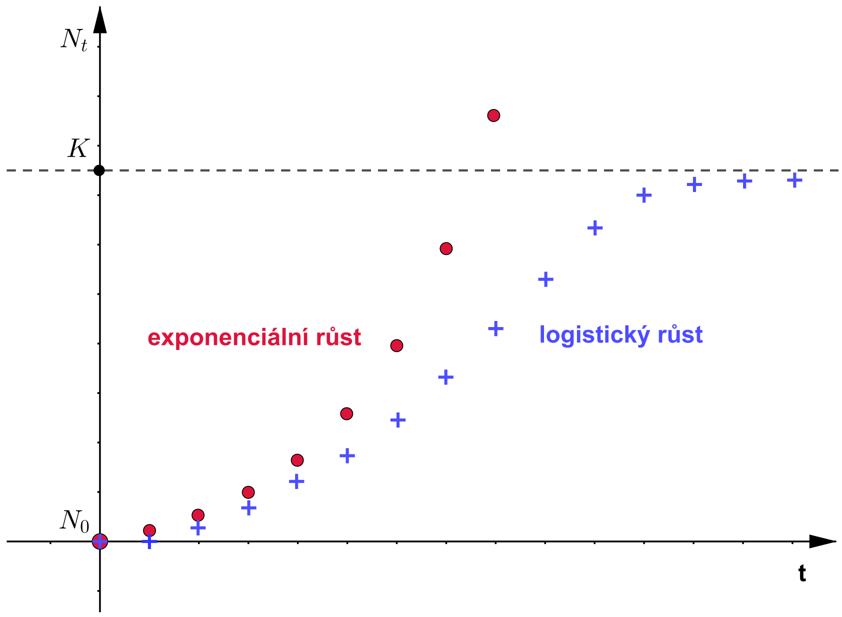
Takovéto limitující negativní vlivy určují tzv. **nosnou kapacitu prostředí** (neboli mez únosnosti), tedy maximální udržitelnou velikost populace, která se může na daném území s dostupnými zdroji uživit, aniž by došlo ke snížení úživnosti této populace v budoucnosti. Nosnou kapacitu prostředí označujeme zpravidla K. [9]

Studie B. Pengra z roku 2012 „One Planet, How Many People?“ shrnula 65 různých odhadů maximální udržitelné velikosti lidské populace. Nejčastější odhad byl okolo 8 miliard, což je číslo, kterého lidská populace dosáhne během 10 let. Mezi další uvedené odhady ovšem patří také čísla jako 2 miliardy nebo 1024 miliard.[10]

Nepočetné a řídké populace mohou růst v prostředí bez kompetice zprvu velmi rychle. Tento růst nazýváme **exponenciální růst populace** (viz obr. 3.1). Jeho růstová křivka je ve tvaru písmene „J“. S dramatickým nárůstem populační hustoty se zvyšují také nároky na zdroje, kterých stále ubývá. Následkem toho se začíná projevovat vnitrodruhová kompetice (tj. soupeření mezi jedinci téhož druhu)[1] a v jejím důsledku se růst populace zpomaluje.

[1]přirozeně dochází také ke kompetici mezi jedinci různého druhu, tu zde ale uvažovat nebudeme

Natalita se přibližuje k mortalitě, a populace tak nabude své nosné kapacity. Efekt kompetice exponenciální růstovou křivku deformuje do tvaru „S“, tento typ růstu populace označujeme jako **logistický růst** (viz obr. 3.1) a jeho růstovou křivku jako S-křivku. [8]



Obrázek 3.1: Exponenciální a logistický růst v diskrétním případě (tj. určitých časových okamžicích). Čas označen t, počáteční velikost populace N_0, velikost populace v čase t označena N_t a nosná kapacita prostředí K.

Exponenciálně mohou růst například populace nového (**invazivního**) druhu zavlečeného nebo úmyslně vysazeného (**introdukovaného**) např. na ostrov; dále populace bakteriálních a virových infekcí (např. mor ve středověku, virus SARS) nebo různé v laboratorních podmínkách pěstované populace. Vždy se jedná o populace, které nejsou nijak limitovány (predace, nemoci,...) a mají dostatek pro život důležitých zdrojů.

Z historie známými příklady exponenciálního růstu jsou:

- expanze severoamerické ondatry (*Ondatra zibethicus*), kterou v roce 1905 vysadil na svém panství v Dobříši u Prahy kníže Mansfeld, původně čítala populace 5 kusů, během 22 let se ovšem rozšířila do celé střední Evropy,
- expanze králíka divokého (*Oryctolagus cuniculus*) introdukovaného po roce 1859 v jižní Austrálii, kde se jeho populace rozšířila o 130 km každý rok,
- populace bažanta (*Phasianus*) introdukovaná v roce 1937 na ostrově Protection

v americkém státě Washington, která vzrostla během pouhých 5 let ze 2 samců a 6 samic na 1325 jedinců,

- populace prasete divokého (*Sus scrofa*) v České republice během posledních 50 let, jejíž roční nárůst je průměrně okolo 11 % (v roce 1950 uloveno 148 jedinců v roce 2004 přes 120 tisíc jedinců!). [9]

Jednou z metod studia populační dynamiky je **matematické modelování**. Matematické modelování je vědecká disciplína umožňující srozumitelně popsat pomocí matematického jazyka chování určitého studovaného systému. Do matematického modelu se úmyslně zahrnují důležité faktory vnější reality a ty méně důležité se zanedbávají. Model umožňuje odhalit podstatné vztahy mezi prvky systému, případně predikovat chování systému v budoucnosti. Mezi nejzákladnější matematické modely populační dynamiky patří **exponenciální model** (nazýván také model populačního růstu nezávislého na hustotě nebo Malthusův model), **logistický model** (nazýván také Verhulstův model), **Rickerův model**, a **Leslieho model** zohledňující na rozdíl od předchozích modelů věkovou strukturu. [9]

Více o zmíněných populačních modelech viz [9] od s. 87.

Shrnutí: Populace tedy roste exponenciálně pouze v pro ni příznivých podmínkách, kdy je všeho dostatek (živin, hostitelů, potravy, prostoru, úkrytů...). Každá populace však musí být něčím omezována, buď právě zdroji, nebo je někým či něčím devastována, např. požírána predátorem, odnášena vodou, spalována požáry, hubena člověkem apod. Exponenciální růst je vždy **dočasný**, dříve či později přestává populace prosperovat a růst se zpomaluje až dosáhne nosné kapacity prostředí.

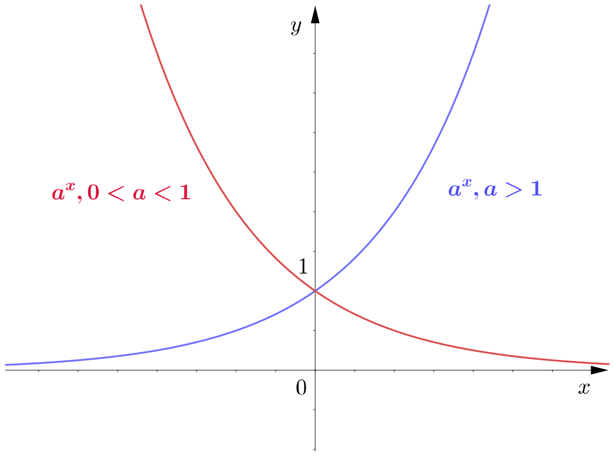
3.2 Exponenciální a logaritmická funkce

Než si přestavíme exponenciální růst populace, zopakujeme si několik nejdůležitějších poznatků o exponenciální funkci a k ní inverzní logaritmické funkci.

Definice 3.2.1. ([11], s. 125)[**Exponenciální funkce.**] Exponenciální funkce o základu a má předpis

$$f(x) = a^x, \tag{3.1}$$

kde $x \in \mathbb{R}$, $a \in \mathbb{R}^+$ a $a \neq 1$.



Obrázek 3.2: Graf exponenciální funkce $y = a^x$ pro rozdílné hodnoty základu a.

Z grafu exponenciální funkce (viz obr. 3.2) lze vyčíst některé důležité vlastnosti této funkce:

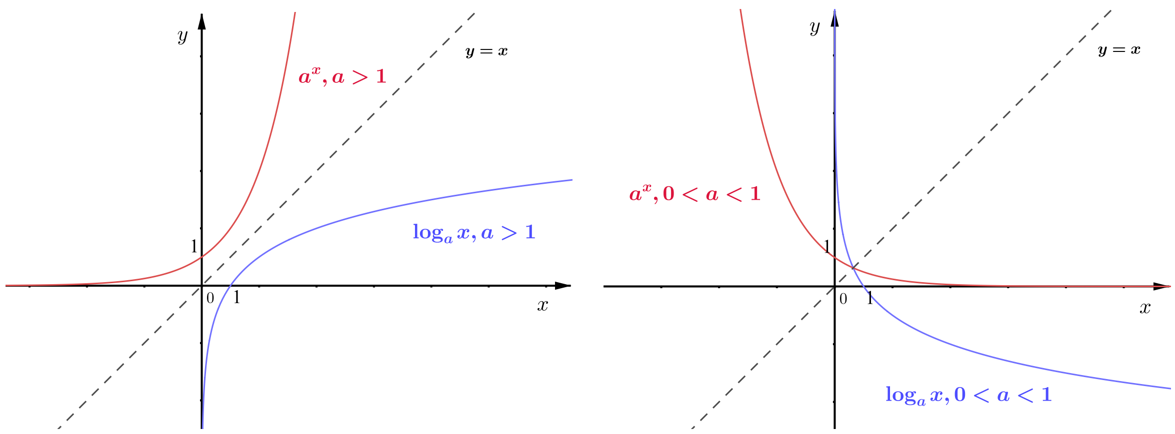
- graf funkce prochází bodem $[0; 1]$, neboť $a^0 = 1$,
- její definiční obor je $D_f = \mathbb{R}$ a obor hodnot $H_f = (0, \infty)$,
- je-li její základ $a > 1$, je rostoucí a je-li $0 < a < 1$, je klesající,
- je prostá a zdola omezená nulou na celém svém definičním oboru.

Definice 3.2.2. ([11], s. 134)[**Logaritmická funkce.**] Logaritmická funkce o základu a je funkce, která je inverzní k exponenciální funkci $f(x) = a^x$ a má tvar

$$f^{-1}(x) = \log_a x, \tag{3.2}$$

kde $x > 0$, $a \in \mathbb{R}^+$ a $a \neq 1$.

Poznámka 3.2.1. Je-li speciálně ve vzorci (3.1) (resp. (3.2)) $a = e$, tj. základ a je roven konstantě $e \approx 2,718281$ nazývané Eulerovo číslo, označujeme exponenciální (resp. logaritmickou) funkci jako přirozená exponenciální (resp. přirozená logaritmická) funkce.



Obrázek 3.3: Graf exponenciální funkce a k ní inverzní logaritmické funkce – vlevo pro $a > 1$, vpravo pro $0 < a < 1$.

Následující vlastnosti logaritmické funkce vyplývají ze skutečnosti, že logaritmická funkce je inverzní k funkci exponenciální (viz obr. 3.3):

- graf funkce prochází bodem $[1; 0]$,
- její definiční obor je $D_{f^{-1}} = H_f = (0, \infty)$, a obor hodnot $H_{f^{-1}} = D_f = \mathbb{R}$,
- je-li její základ $a > 1$, je rostoucí a je-li $0 < a < 1$, je klesající,
- je prostá a není omezená.

Za připomenutí stojí také skutečnost, že grafy funkce a k ní příslušné inverzní funkce jsou vždy souměrné podle osy $y = x$ (viz obr. 3.3).

Definice 3.2.3. ([11], s. 139)[**Logaritmus.**] Logaritmus čísla y o základu a je takové číslo x, pro které platí $a^x = y$:

$$\log_a y = x \text{ , právě když } a^x = y.$$

Platí následující vztahy:

1. $\log_a a = 1, \ \log_a 1 = 0,$
2. $\log_a x^n = n \cdot \log_a x, \ n \in \mathbb{R}$
3. $\log_a (xy) = \log_a x + \log_a y, \ \ \log_a \left(\frac{x}{y}\right) = \log_a x - \log_a y.$

Chceme-li formálně najít inverzní funkci k exponenciální funkci, musíme vlastně vyjádřit x, což uděláme tak, že exponenciální funkci zlogaritmujeme a upravíme podle výše uvedených vztahů:

$$\begin{aligned} y &= a^x \qquad /\log_a \\ \log_a y &= \log_a a^x \\ \log_a y &= x \cdot \log_a a \\ \log_a y &= x \end{aligned}$$

Příklad 3.2.1. Určete exponent x, je-li $6^x = 216$.

$$\begin{aligned} 216 &= 6^x \qquad /\log_6 \\ \log_6 216 &= \log_6 6^x \\ \log_6 216 &= x\log_6 6 \\ \log_6 216 &= x \ \Rightarrow \ x = 3, \text{ protože } 6^3 = 216 \end{aligned}$$

3.3 Exponenciálně rostoucí populace

Model populace s nepřekrývajícími se generacemi

Nejprve uvažujme **populaci s nepřekrývajícími se generacemi množící se pouze v určité sezóně**, typicky některé druhy hmyzu nebo jednoleté rostliny. Jako příklad uveďme motýly, kteří v létě nakladou vajíčka a dospělci po rozmnožení rychle umírají. Z vajíček se poté vylíhnou housenky, které se na podzim zakuklí. Kukly přežijí zimu, na jaře se z nich vylíhnou dospělí motýli a cyklus se opakuje.

Pokud se jedná o jedince s odděleným pohlavím, zahrnujeme do modelu pouze počty samic a samičích mláďat (samce zanedbáváme).

Označme:

t ... čas, typicky sezóna (dále např. rok, den, hodina),

N_0 ... počet jedinců v nulté sezóně (počáteční velikost populace),

N_t ... velikost populace v sezóně t,

N_{t+1} ... velikost populace v příští sezóně

α ... tzv. **konečná růstová rychlost**, která udává, kolik dospělý jedinec zplodí nových jedinců, kteří se dožijí příští sezóny a plodí.[9]

Sledujeme-li počet dospělců každý rok, potom počet jedinců první v sezóně bude

$$N_1 = N_0\alpha,$$

v druhé sezóně pak

$$N_2 = N_1\alpha = N_0\alpha^2,$$

atd. Obecně po t sezónách získáme vztah

$$N_t = N_0\alpha^t. \tag{3.3}$$

Je zřejmé, že α je konstanta a není závislá na čase, velikosti populace, ani jiných faktorech. Obecný vztah pro počet jedinců v sezóně $t+1$ je tvaru

$$N_{t+1} = N_t\alpha. \tag{3.4}$$

Z předchozího vzorce vidíme, že závislost populační velikosti N na čase je exponenciální. Přírůstek α za danou sezónu získáme snadno jako

$$\alpha = \frac{N_{t+1}}{N_t}. \tag{3.5}$$

Tento typ diskrétního populačního růstu (jedinci se množí pouze v určité sezóně) se také označuje jako **geometrický**.[9]

Je zřejmé, že rovnice (3.4) popisuje geometrickou posloupnost $\{N_t\}$ s kvocientem α a (3.5) je vzorec pro výpočet kvocientu ze dvou po sobě jdoucích členů geometrické posloupnosti.

Pokud je $\alpha > 1$, řešení (3.3) vede na **exponenciální růst**, pokud je naopak $0 < \alpha < 1$, jedná se o **exponenciální vymírání**. Pro $\alpha = 1$ zůstává počet jedinců v populaci konstantní.

Příklad 3.3.1. Jednoletka osívka jarní (*Erophila verna*) je drobná, bíle kvetoucí jarní bylinka. Množí se výhradně semeny, kterých jedna rostlina může vyprodukovat až 1000. Mohou klíčit od jara až do podzimu, ale dokáží i několik let ležet v půdě a čekat na vhodný okamžik.[2]

Uvažujte hypotetickou situaci: Jedna osívka dá vznik 20 semenům, ze kterých další rok vyrostou nové rozmnožování schopné rostliny (ostatní semena vždy zahynou např. následkem extrémně studených zim). Neboť se jedná o jednoleté rostliny, dospělé osívky vždy před

[2]informace převzaty z: https://cs.wikipedia.org/wiki/Osívka_jarní

zimou zahynou, takže nedochází k překryvu generací. V populaci je v nulté sezóně 100 jedinců. Předpokládejte exponenciální růst.

1. Určete konečnou růstovou rychlost α.

2. Vypočítejte, kolik jedinců v populaci bude po roce, po dvou letech a po pěti letech.

Řešení:

1. Dle uvedených údajů je $\alpha = 20$, protože jedna osívka za sezónu vyprodukuje 20 jedinců v dalším roce schopných rozmnožování.

2. Po roce bude v populaci $N_1 = 20N_0 = 20 \cdot 100 = 2\,000$ jedinců, po dvou letech to bude $N_2 = 20^2 N_0 = 400 \cdot 100 = 40\,000$ jedinců. Po pěti letech bude mít populace osívky $N_5 = 20^5 N_0 = 20^5 \cdot 100 = 3,2 \cdot 10^8$ jedinců.

Příklad 3.3.2. Martináč hrušňový (*Saturnia pyri*) je největší evropský motýl vyskytující se mimo jiné také na jižní Moravě. Dospělí jedinci mají zakrnělý sosák, a proto nemůžou přijímat potravu. Život dospělců trvající jen několik dní má jediné poslání – rozmnožit se.[3] Každá samička martináče dané hypotetické populace naklade v sezóně mnoho vajíček, ale pouze α vajíček přežije do další sezóny, kdy se opět rozmnoží. Dospělec po nakladení vajíček rychle umírá. Opět předpokládejte exponenciální růst.

1. Určete koeficient α.

2. Doplňte chybějící údaje v následující tabulce (N označuje počty samic v daných sezónách).

N_0	N_1	N_2	N_3	N_4	...	
		17 150		840 350	...	288 240 050

[3] informace převzaty z: https://cs.wikipedia.org/wiki/Martinačovití

Řešení:

1. Koeficient α určíme z velikosti populací N_2 a N_4 zanesených v tabulce:

$$N_4 = N_2\alpha^2 \;\Rightarrow\; \alpha = \sqrt{\frac{N_4}{N_2}} = \sqrt{\frac{840\,350}{17\,150}} = \sqrt{49} = 7.$$

Zjistili jsme, že $\alpha = 7$.

2. Nyní již snadno vypočítáme chybějící údaje v tabulce:

$$\begin{aligned} N_2 &= N_0\alpha^2 \;\Rightarrow\; N_0 = \frac{N_2}{\alpha^2} = \frac{17\,150}{49} = 350, \\ N_1 &= N_0\alpha = 350 \cdot 7 = 2\,450, \\ N_3 &= N_0\alpha^3 = 350 \cdot 7^3 = 120\,050. \end{aligned}$$

Jako poslední určíme, za kolik sezón bude mít populace martináče 288 240 050 jedinců. Uplatníme postup z Příkladu 3.2.1, kde jsme také zjišťovali příslušný exponent. Obě strany rovnice $N_t = N_0\alpha^t$ nejprve vydělíme N_0, poté zlogaritmujeme a upravíme následovně:

$$\begin{aligned} \frac{N_t}{N_0} &= \alpha^t \\ \frac{288\,240\,050}{350} &= 7^t \\ 823\,543 &= 7^t \\ \log_7 823\,543 &= t \\ t &= 7 \end{aligned}$$

Populace dosáhne velikosti 288 240 050 jedinců za 7 sezón. Vyplněná tabulka tedy vypadá následovně:

N_0	N_1	N_2	N_3	N_4	...	$\mathbf{N_7}$
350	**2 450**	17 150	**120 050**	840 350	...	288 240 050

Model populace s překrývajícími se generacemi a s jednorázovým rozmnožováním

Nyní uvažujme **populace s překrývajícími se generacemi a s jednorázovým rozmnožováním**. Uvažujeme v tomto případě **pouze samice** a jejich sčítání provádíme vždy těsně před tím, než samice rodí. Příkladem mohou být zajíci množící se v dubnu. Změna velikosti populace je ovlivněna pouze počtem nově narozených a zemřelých samic, ostatní vlivy zanedbáme.

Označme:

N_t ... počet samic zjištěný těsně před rozmnožením v čase t,

$n \geq 0$ ... koeficient natality (počet samičích mláďat na jednu dospělou plodnou samici, kteří se dožijí příštího jara – tj. průměrný počet narozených mláďat $\times$ pravděpodobnost jejich přežití),

$m \geq 0$ ... koeficient mortality (procento dospělých plodných samic, které nepřežijí z jednoho jara do druhého)

Platí tedy, že počet samic v následujícím roce bude rovný počtu dospělých samic v přítomném roce spolu s nově narozenými životaschopnými samicemi a bez dospělých samic, které nepřežijí do následujícího roku:[12]

$$N_{t+1} = N_t + nN_t - mN_t = N_t(1 + n - m).$$

Označme $1 + n - m = \alpha$, čímž získáme opět klasickou, nám již známou, rovnici exponenciálního růstu s růstovým koeficientem α :

$$N_{t+1} = N_t(1 + n - m) = N_t\alpha. \tag{3.6}$$

$$N_t = N_0\alpha^t.$$

Jestliže $n > m$ (tj. podíl nově narozených převyšuje podíl zemřelých jedinců za jednotku času), populace exponenciálně **roste**, je-li naopak $n < m$, pak populace exponenciálně **vymírá**. Pro $n = m$ je velikost populace konstantní.

Příklad 3.3.3. Zajíc polní (*Lepus europaeus*) žije mimo období páření samotářsky a obhajuje si své teritorium. V době páření (honcování), které začíná v únoru a trvá až do srpna nebo září, se zajíci sdružují v malé hloučky. Zpozorujeme-li na poli za sebou pobíhat několik zajíců, první z nich je říjná zaječka, kterou pronásleduje skupina zajíců. Rodí se obvykle 2 – 5 mláďat a to několikrát do roka, obvykle však do druhého roku přežívají pouze 2 – 4 mláďata.[4]

Uvažujte hypotetickou expanzivní populaci zajíců, kterým se podaří rozmnožit pouze jednou do roka.

1. Každé zaječici se narodí v jedné sezóně v průměru 3 malé zaječice a úmrtnost samic během roku je 55 %. V sezóně 0 čítá populace 25 zaječic. Rozhodněte, jestli populace

[4]informace převzaty z: https://cs.wikipedia.org/wiki/Zajíc_polní

exponenciálně roste, nebo exponenciálně vymírá a spočítejte, jaká bude velikost populace za 3 roky.

2. Nyní uvažujte možnost, že v dané populaci následně propukla tularémie – bakteriální onemocnění zajíců a hlodavců přenášené klíšťaty. Porodnost samic po třetím roce klesla na 0,5 mláďata za rok a úmrtnost se zvyšuje na $85\ \%$. Rozhodněte, jestli populace roste, nebo exponenciálně vymírá a spočítejte, jaká bude velikost populace na konci následujícího (tj. čtvrtého) roku.

Řešení:

1. Nejprve určíme koeficient $\alpha_1 = 1 + 3 - 0,55 = 3,45$. Protože $\alpha_1 > 1$, populace zajíců exponenciálně roste. Nyní zjistíme velikost populace za 3 roky: $N_3 = N_0\alpha_1^3 = 25 \cdot 3,45^3 \cong 1\ 027$.

2. Nejprve opět určíme koeficient $\alpha_2 = 1 + 0,5 - 0,85 = 0,65$. Protože $0 < \alpha_2 < 1$, populace exponenciálně vymírá. Nyní zjistíme velikost populace nakažené tularémií následující rok: $N_4 = N_3\alpha_2 = 1\ 027 \cdot 0,65 \cong 668$. Tularémií infikovaná populace bude na konci čtvrtého roku čítat jen 668 samic.

Model populace s překrývajícími se generacemi a kontinuálním přírůstkem

Nakonec uvažujme **populace s překrývajícími se generacemi a s kontinuálním přírůstkem**. Růst je sice také exponenciální, ale odvození je v tomto případě složitější (během jedné sezóny se samice kontinuálně rodí i umírají; samice, které umřely na začátku časového úseku, potom už nerodí).

Tato situace se řeší tak, že se zkracuje uvažovaný časový úsek až do podoby **diferenciální rovnice:**[5]

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = nN(t) - mN(t), \tag{3.7}$$

kde $N(t)$ je nám neznámá funkce představující počet jedinců populace (velikost populace) v čase t, n je natalita a m je mortalita. Zlomek $\dfrac{\mathrm{d}N(t)}{\mathrm{d}t}$ značí derivaci funkce $N(t)$ podle

[5] Diferenciální rovnice jsou rovnice obsahující neznámou funkci jedné nezávislé proměnné a její derivace.

proměnné t, tj. $N'(t)$, a vyjadřuje změnu velikosti populace v čase. Položíme-li v rovnici (3.7) $r = n - m$, obdržíme diferenciální rovnici

$$N'(t) = rN(t), \tag{3.8}$$

kde r je **(relativní) velikost změny**.[9]

Opět platí, že je-li $n > m$, populace exponenciálně roste, je-li naopak $n < m$, pak populace exponenciálně vymírá, a pro $n = m$ je velikost populace v čase neměnná.

Diferenciální rovnici (3.8) vyřešíme tzv. metodou separace proměnných, kdy nejdříve převedeme funkce nezávislé proměnné na jednu stranu a závislé proměnné na druhou stranu rovnice, a následně obě strany rovnice zintegrujeme.[6]

$$\begin{aligned}
N'(t) &= rN(t) \qquad / : N(t) \\
\frac{N'(t)}{N(t)} &= r \qquad / \int \\
\int \frac{\mathrm{d}N(t)}{\mathrm{d}tN(t)}\mathrm{d}t &= \int r \,\mathrm{d}t \\
\int \frac{1}{N(t)}\mathrm{d}N(t) &= r \int \,\mathrm{d}t \\
\ln|N(t)| &= rt + K \qquad /e \\
|N(t)| &= e^{rt+K} \\
N(t) &= e^{rt} \cdot N(0),
\end{aligned}$$

kde $K, N(0)$ jsou libovolné reálné konstanty. Správnost řešení lze ověřit jeho zderivováním: $N'(t) = (N(0)e^{rt})' = rN(0)e^{rt} = rN(t)$.

Řešením rovnice (3.8) je tedy opět rovnice exponenciálního růstu

$$N(t) = N(0) \cdot e^{rt}, \tag{3.9}$$

kde $r = n - m$ a $N(0)$ je počáteční velikost populace, kterou musíme nezbytně znát, abychom mohli určit velikost populace $N(t)$. Zlogaritmováním obou stran rovnice (3.9) získáme vztah pro výpočet růstové rychlosti r:[9]

$$r = \frac{\ln N(t) - \ln N(0)}{t} = \frac{\ln \frac{N(t)}{N(0)}}{t}.$$

[6]Na středních školách se studenti obvykle nesetkají s diferenciálními rovnicemi a jejich řešením, proto nemá smysl se zde jimi detailněji zabývat. Vyřešení rovnice (3.8) pomocí integrace lze však zadat matematicky nadaným studentům např. jako samostudium.

Příklad 3.3.4. V Indii byla v roce 2017 natalita 19 narození na 1 000 obyvatel a mortalita 7,3 úmrtí na 1 000 obyvatel. Velikost populace Indie byla v témže roce 1 281 935 911.[7]

1. Určete růstovou rychlost dané populace (předpokládejte kontinuální přírůstek).

2. Za jak dlouho dosáhne populace Indie 1,5 miliardy (při zachování konstantní růstové rychlosti)?

3. V jakém roce se populace Indie zdvojnásobí (při zachování konstantní růstové rychlosti)?

Řešení:

1. Nejprve je nutné si vyjádřit natalitu a mortalitu v promile: $n = 0,019$ rok^{-1} a $m = 0,0073$ rok^{-1}. Nyní už můžeme vypočítat růstovou rychlost: $r = n - m = = 0,019 - 0,0073 = 0,0117$.

2. Známe počáteční velikost populace $N(0) \cong 1,28$ miliardy, růstovou rychlost $r = 0,0117$ a výslednou velikost $N(t) = 1,5$ v neznámém čase, který máme zjistit. K výpočtu použijeme vzorec $N(t) = N(0)e^{rt}$. Obě strany této rovnice vydělíme $N(0)$ a postupujeme následovně:

$$\begin{aligned}\frac{N(t)}{N(0)} &= e^{rt}\\ \frac{1,5}{1,28} &= e^{0,0117t} \qquad /\ln\\ \ln\frac{1,5}{1,28} &= 0,0117t\\ t &\cong 13,6.\end{aligned}$$

   Populace Indie dosáhne velikosti 1,5 miliardy přibližně za 13 let a 7 měsíců.

3. Známe počáteční velikost populace $N(0) \cong 1,28$, růstovou rychlost $r = 0,0117$ a výslednou velikost $N(t) = 2N(0) = 2,56$ miliardy v neznámém čase, který máme zjistit. Opět použijeme vzorec $N(t) = N(0)e^{rt}$, do kterého dosadíme $N(t) = 2N(0)$,

[7] data převzata z: https://www.indexmundi.com/india/demographics_profile.html

a upravujeme následovně:

$$\begin{aligned} 2N(0) &= N(0)e^{rt} \qquad / : N(0) \\ 2 &= e^{0,0117t} \qquad / \ln \\ \ln 2 &= 0,0117t \\ t &= \frac{\ln 2}{0,0117} \cong 59,24 \end{aligned}$$

Populace Indie dosáhne dvojnásobné velikosti 1,5 miliardy přibližně v roce 2076 (vycházíme z počátečního roku 2017).

Nosná kapacita prostředí

V podkapitole 3.1 Motivace již byl vysvětlen pojem **nosná kapacita prostředí** K. Nosná kapacita je vlastně jeden konkrétní počet jedinců K dané populace, ke kterému se od určitého mezního okamžiku t blíží (popř. kolem kterého kolísají) všechny velikosti populace $N_{t+1}, N_{t+2}, N_{t+3}, ...$, a s časem se tato hodnota K nemění. Názorně viz obr. 3.1. Nosnou kapacitu lze tedy přibližně učit jako aritmetický průměr z těchto hodnot $N_{t+1}, N_{t+2}, N_{t+3}, ..., N_{t+n}$.

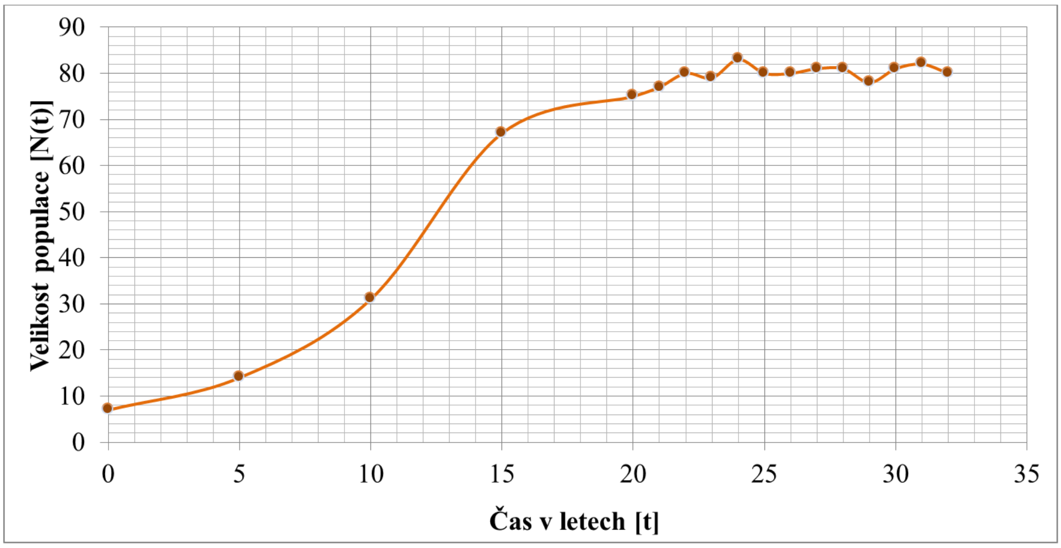
Příklad 3.3.5. Na jistý malý ostrůvek byl jako „černý pasažér“ obchodní lodi člověkem nevědomě zavlečen kontinuálně se množící druh hlodavce – krysa obecná (*Rattus rattus*). Počáteční velikost populace byla 7 jedinců. Tomuto invazivnímu druhu se na ostrově, kde měl málo přirozených nepřátel a spoustu prostoru, velmi dařilo, proto počet jedinců jeho druhu rychle stoupal. Zjištěné velikosti populace v jednotlivých letech jsou zaneseny v následující tabulce a grafu:[8]

t	0	5	10	15	20	21	22	23	24	25	26	27	28	29	30	31	32
$N(t)$	7	14	31	67	75	77	80	79	83	80	80	79	81	82	81	78	80

1. Pokuste se z hodnot v tabulce a pomocí grafu (obr. 3.4) odhadnout nosnou kapacitu ostrova K pro danou krysí populaci.

2. Předpokládejte, že během prvních patnácti let rostla populace exponenciálně, na základě tohoto předpokladu určete přibližnou rychlost růstu r v prvních 15 letech.

[8] jedná se o vlastní příklad



Obrázek 3.4: Graf ukazující závislost velikosti populace invazivní krysy na čase.

Řešení:

1. Z tabulky i grafu je patrné, že přibližně od 21. roku hodnoty $N(t)$ kolísají mezi hodnotami 77 a 83. Nosnou kapacitu K tedy určíme přibližně jako aritmetický průměr z hodnot $N(21), N(22), ..., N(32)$:
$$K = \frac{77 + 80 + 79 + 83 + 80 + 80 + 79 + 81 + 82 + 81 + 78 + 80}{12} = 80.$$
V tomto konkrétním případě je K přibližně 80 jedinců.

2. Z údajů v tabulce víme, že se z počáteční velikosti populace zvýšila velikost populace za 5 let na 14 jedinců, za 10 let už měla 31 jedinců a po 15 letech 67 jedinců. Hodnoty postupně zadáme do vzorce (3.9) a zlogaritmováním obou stran vypočítáme koeficienty r:
$$N(5) = N(0)e^{5r_1} \;\Rightarrow r_1 = \frac{\ln \frac{N(5)}{N(0)}}{5} = \frac{\ln \frac{14}{7}}{5} \cong 0,138,$$
$$N(10) = N(0)e^{10r_2} \;\Rightarrow r_2 = \frac{\ln \frac{N(10)}{N(0)}}{10} = \frac{\ln \frac{31}{7}}{5} \cong 0,149,$$
$$N(15) = N(0)e^{15r_3} \;\Rightarrow r_3 = \frac{\ln \frac{N(15)}{N(0)}}{15} = \frac{\ln \frac{67}{7}}{15} \cong 0,151,$$

Růstová rychlost dané populace byla v prvních 15 letech přibližně mezi 0,14 až 0,15.

3.4 Pracovní list

Téma: Aplikace exponenciální funkce v populační ekologii

Jméno a třída: Datum:

Příklad č. 1 Studujeme invazivní druh slunéčka východního (*Harmonia axyridis*) zavlečeného z východní Asie do Evropy. Slunéčka se reprodukují pouze jednou ročně a hynou ihned po nakladení vajíček. Z vajíček se vylíhne larva, která se mění v kuklu, ta poté přečká zimu.[9] Z vylíhnutých vajíček nakladených jednou samičkou přežije do další sezóny průměrně 15 nových reprodukce schopných samiček. Na jednom hostitelském stromě, kde se slunéčka živí mšicemi, bylo na počátku měření zjištěno 6 samiček slunéček. Zanedbejte vliv nemocí i predátorů na velikost populace a předpokládejte exponenciální růst.

1. Na základě zadání rozhodněte, o jaký model exponenciálně rostoucí populace se jedná (generace se překrývají / nepřekrývají; rozmnožování je jednorázové / kontinuální). Své rozhodnutí zdůvodněte.

2. Určete růstovou rychlost populace slunéček a vypočítejte, kolik bude mít tato populace samiček příští rok a za 4 roky.

[9] informace převzaty z: https://cs.wikipedia.org/wiki/Slunéčko_východní

Příklad č. 2 Africký stát Mali se řadí k lidským populacím s největší natalitou na světě, v roce 2017 to bylo 43,9 na 1 000 obyvatel. Mortalita v témže roce byla 9,8 na 1 000 obyvatel a počet obyvatel přibližně 17,9 milionů.[10]

1. Na základě zadání rozhodněte, o jaký model populace s exponenciálním růstem se jedná (generace se překrývají / nepřekrývají; rozmnožování je jednorázové / kontinuální). Své rozhodnutí zdůvodněte.

2. Určete růstovou rychlost populace Mali a vypočítejte, za kolik let se velikost populace ztrojnásobí.

[10] data převzata z: https://www.indexmundi.com/mali/demographics_profile.html

Příklad č. 3 Mějme hypotetickou populaci prvoka trepky velké (*Paramecium caudatum*) pěstované v laboratorních podmínkách, kde se rozmnoží pouze nepohlavně příčným dělením jednou za 8 hodin. Dospělý jedinec tedy žije i po rozmnožení. Trepka velká se běžně nachází v organicky znečištěných vodách celého světa. Jedná se o největší z rodu trepek o délce 0,17 – 0,35 mm a je pozorovatelná i pouhým okem.[11]

1. Předpokládejte exponenciální růst. Na základě zadání rozhodněte, o jaký model exponenciálně rostoucí populace se jedná (generace se překrývají / nepřekrývají; rozmnožování je jednorázové / kontinuální). Své rozhodnutí zdůvodněte.

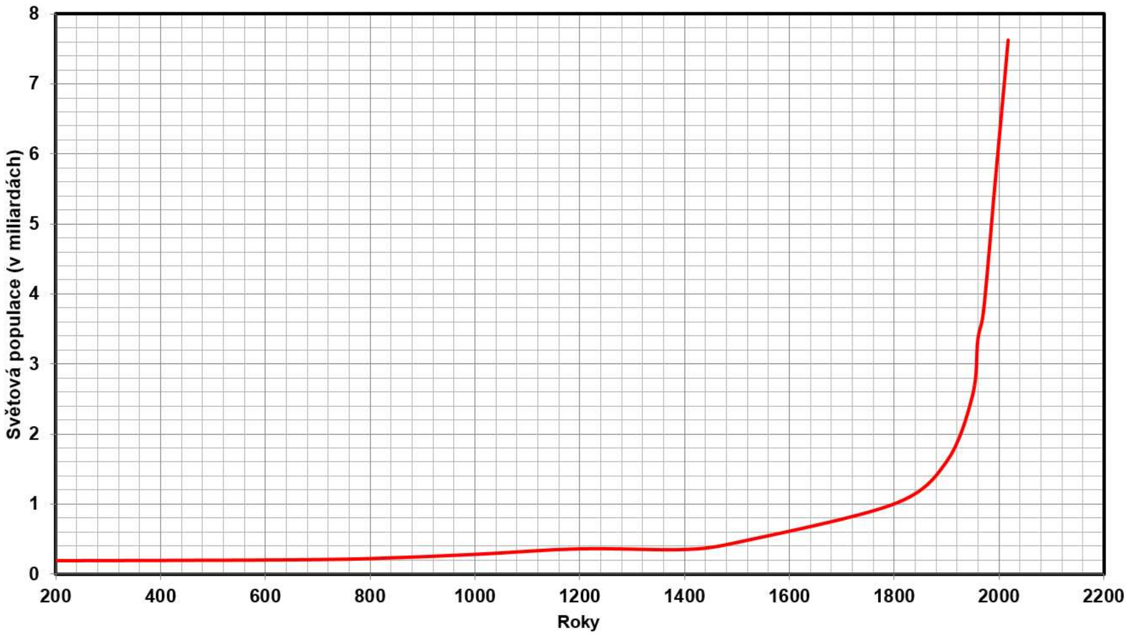
2. Určete růstovou rychlost populace. Doplňte chybějící údaje v tabulce, které zapomněl doplnit vědec prof. Emil Roztržitý provádějící experiment s laboratorně pěstovanými trepkami:

hodiny	0	8	16	24	48	72
počet jedinců		9			170	

[11] informace převzaty z: https://cs.wikipedia.org/wiki/Trepka_velká

Bonusový příklad: V následující tabulce a grafu jsou zaneseny údaje o počtu jedinců lidské světové populace v letech 200 – 2018 (v miliardách):[12]

rok	**200**	**600**	**800**	**1000**	**1200**	**1400**	**1500**	**1800**	**1900**
populace	0,19	0,2	0,22	0,28	0,36	0,35	0,45	1	1,6
rok	**1950**	**1960**	**1970**	**1980**	**1990**	**2000**	**2010**	**2018**	
populace	2,55	3,34	3,7	4,46	5,33	6,15	6,96	7,63	



Obrázek 3.5: Graf vývoje světové populace v letech 200 – 2018 (v miliardách), vytvořen dle údajů z předchozí tabulky.

1. Na základě údajů v tabulce a grafu vývoje světové populace rozhodněte, zda je možné pokládat růst světové populace za exponenciální. Své rozhodnutí zdůvodněte.
2. Předpokládejte nyní (nezávisle na předchozích zjištěních), že lidská světová populace roste v posledních letech exponenciálně. Pokuste se odhadnout, kolik lidí by v tomto případě bylo na světě v roce 2020, 2050 a 2100. K výpočtu růstové rychlosti použijte data z let 2010 a 2018.

[12]data převzány z: http://www.worldometers.info/world-population/#pastfuture

Řešení pracovního listu:

Příklad č. 1:

1. Model s nepřekrývajícími se generacemi a jednorázovým rozmnožením (dospělci po vykladení umírají a reprodukují se jednou za rok).

2. Ze zadání víme, že růstová rychlost α je 15 (protože do další sezóny přežije právě 15 nových reprodukce schopných samiček). Příští rok bude mít populace $N_1 = 6 \cdot 15 = 90$ samiček a za 4 roky $N_4 = 6 \cdot 15^4 = 303\ 750$ samiček.

Příklad č. 2:

1. Model s překrývajícími se generacemi a kontinuálním rozmnožováním (lidé se rodí po celý rok a ženy přirozeně ihned po porodu neumírají).

2. Růstová rychlost je $r = n - m = 0,0439 - 0,0098 = 0,0341$ (natalitu a mortalitu je třeba si vyjádřit v promile!). K dalšímu výpočtu použijeme vzorec $N(t) = N(0)e^{rt}$. Ze zadání víme, že platí $N(0)e^{0,0341t} = 3N(0) \Rightarrow e^{0,0341t} = 3 \Rightarrow t \cong 32,2$. Populace Mali se ztrojnásobí přibližně za 32 let.

Příklad č. 3:

1. Model s překrývajícími se generacemi a jednorázovým rozmnožením (po rozmnožení trepka zpravidla neumírá a množí se jednou za určitou stále stejnou dobu, konkrétně po 8 hodinách).

2. Růstová rychlost je $\alpha^5 = \frac{N_5}{N_0} \Rightarrow \alpha = \sqrt[5]{\frac{170}{9}} \cong 1,8$. Při výpočtu chybějících hodnot v tabulce dosazujeme do vzorce $N_t = N_0\alpha^t$:

hodiny	0	8	16	24	48	72
počet jedinců	5	9	16	29	170	992

Bonusový příklad:

1. Obecně platí, že populace roste exponenciálně, pokud růst není omezen negativní závislostí na hustotě populace (tj. počtem jedinců na určitou jednotku plochy). V důsledku toho jsou parametry růstu (natalita, mortalita, růstová rychlost) konstantní. Není-li tedy daná populace omezována ani nedostatkem zdrojů, ani predátory nebo katastrofami, lze její růst pokládat za exponenciální. Daří-li se naopak populaci konstantně stále špatně bez ohledu na hustotu populace, dochází k exponenciálnímu poklesu.

   Podívejme se nyní na údaje o lidské populaci v zadání – od počátku novověku zřejmě natalita převažuje nad mortalitou a velikost populace neustále roste. Avšak je nutné podotknout, že růstová rychlost není v průběhu lidské historie konstantní, a dokonce v průběhu posledních desítek let dochází k jejímu postupnému zpomalování. Proto růst lidské populace z dlouhodobého hlediska nevyhovuje exponenciálnímu růstu. Pokud ale budeme chtít modelovat růst lidské populace a predikovat jeho vývoj v budoucnu, můžeme růst lidské populace v posledních letech považovat za exponenciální.

2. Budeme předpokládat, že populace roste a bude růst exponenciálně s přibližnou růstovou rychlostí
$$r = \frac{\ln \frac{N(2018)}{N(2010)}}{8} = \frac{\ln \frac{7,63}{6,96}}{8} \cong 0,0115,$$
   což je roční růstová rychlost v posledních 8 letech. Použijeme nyní rovnici $N(t) = N(0)e^{rt}$. Za předpokladu zachování růstové rychlosti $r = 0,0115$ by v roce 2020 bylo na Zemi přibližně
$$N(2020) = 7,63e^{0,0115 \cdot 2} \cong 7,81$$
   miliard lidí, v roce 2050 přibližně
$$N(2050) = 7,63e^{0,0115 \cdot 32} \cong 11,02$$
   miliard lidí a v roce 2100 by to pak bylo přibližně
$$N(2100) = 7,63e^{0,0115 \cdot 82} \cong 19,59$$
   miliard lidí .

   Jedná se však pouze o hrubé odhady, neboť jak již bylo zmíněno, růst lidské populace není v průběhu let konstantní a navíc nebylo zatím přesně určeno, jaká je nosná kapacita prostředí a kdy jí naše populace dosáhne.

3.5 Realizovaná výuka

Dne 14. 6. 2018 jsem odučila na Gymnáziu Jírovcova v Českých Budějovicích v 2. ročníku čtyřletého cyklu vyučovací hodinu na téma Aplikace exponenciální funkce v populační ekologii. Výuka byla pojata jako opakování exponenciální a logaritmické funkce a také jako obohacení učiva o zajímavou aplikaci. Ve třídě bylo v době mé výuky přítomno 29 studentů, průměr známek studentů z matematiky na posledním vysvědčení byl $2,41$.

Studenti byli stejně jako u předchozí odučené kapitoly požádáni o vyplnění krátkého anonymního dotazníku zjišťujícího jejich vztah k matematice a souvislost známek a obliby matematiky, jejich názor na odučenou hodinu a také na povinnou maturitu z matematiky aj. Dotazník se skládal ze dvou částí – první část studenti vyplnili před začátkem výuky, druhou po jejím skončení, a vypadal následovně:

1. část

1. **Baví tě matematika, popř. některá její část?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE
2. **Jakou jsi měl(a) známku z matematiky na posledním vysvědčení?**

   1 – 2 – 3 – 4 – 5
3. **Počítáte v matematice příklady na využití matematiky ve fyzice, technice, ... ?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE
4. **Uvítal(a) bys více příkladů na využití matematiky v přírodních vědách nebo v technice?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE

2. část

1. **Přišla ti vyučovací hodina na téma Aplikace exponenciální funkce v populační ekologii zajímavá?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE
2. **Pomohl ti výklad a příklady v této hodině pochopit exponenciální funkci?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE
3. **Uvítal(a) bys více takových hodin, které by se věnovaly využití matematiky v přírodních vědách?**

   ANO – SPÍŠE ANO – SPÍŠE NE – NE
4. **Jaký je tvůj názor na povinnou maturitu z matematiky?** (otevřená odpověď)

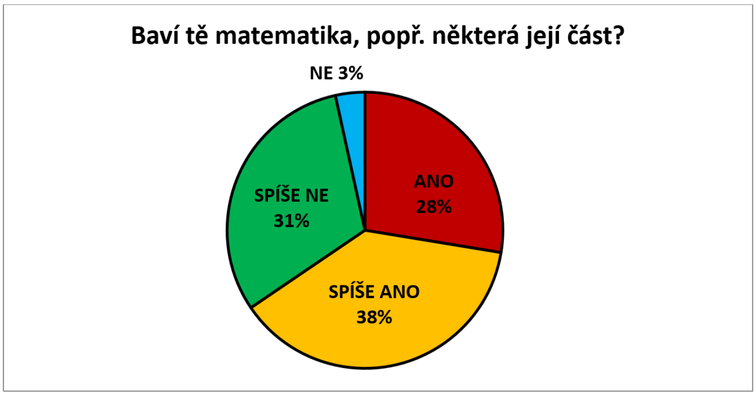
Osnova výuky byla následující:

- vyplnění 1. části dotazníku, představení cílů mojí diplomové práce a výuky,

- vysvětlení základních pojmů: populace, populační ekologie, populační dynamika, zdroje, kompetice o zdroje, nosná kapacita prostředí, exponenciální a logistický růst populace,
- stručné zopakování exponenciální a logaritmické funkce, kdy je daná funkce rostoucí a kdy klesající,
- představení dvou jednoduchých populačních modelů s exponenciálním růstem (model s nepřekrývajícími se generacemi a jednorázovým rozmnožením a model s překrývajícími se generacemi a kontinuálním rozmnožováním), vysvětlení značení, základních vlastností a předvedení příkladu ke každému modelu (příklady 3.3.2 a 3.3.4),
- diskuze na téma růstu lidské populace, porodnosti v rozvojových a rozvinutých zemích, vyplnění 2. části dotazníku a diskuze na téma povinná maturita z matematiky (vztahující se k poslední otázce z dotazníku),
- studentům byl dodatečně poskytnut upravený pracovní list s řešením.

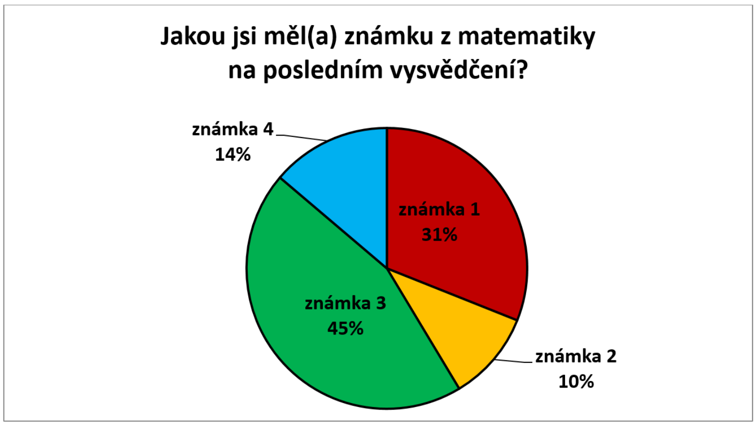
Zpracovaná data z dotazníku

První otázka zjišťovala oblíbenost matematiky u studentů. Ukázalo se, že celkem 66 % studentů baví nebo spíše baví matematika, což je shodou okolností stejná hodnota jako u téže otázky v dotazníku k předchozí kapitole. Opět mne zjištění mile překvapilo.

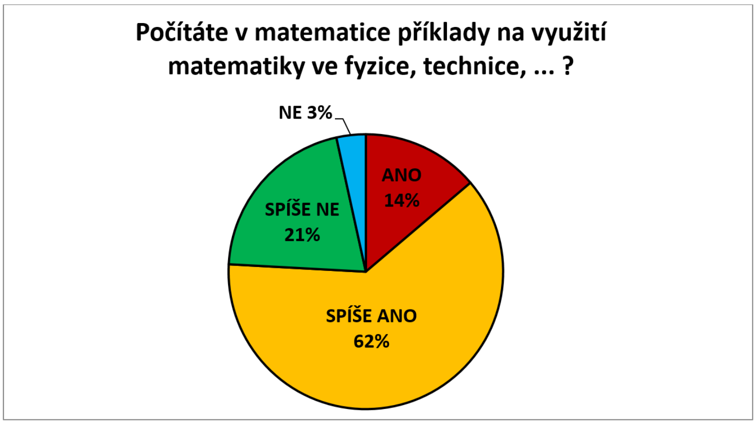


Odpovědi studentů v tomto a předchozím dotazníku se naopak velmi lišily u druhé otázky týkající se jejich známky z matematiky na posledním vysvědčení. Nejčetnější známka ve třídě je 3 s výskytem 45 %, na druhém místě poté 1 s 31 %, následuje 4 s 14 %, nejméně bylo dvojek – pouze 10 %, což odpovídá 3 studentům. Průměr známek ve třídě činil 2,41. Všichni studenti, kteří dostali známku 4, odpověděli, že je matematika nebaví nebo spíše nebaví, u „trojkařů“

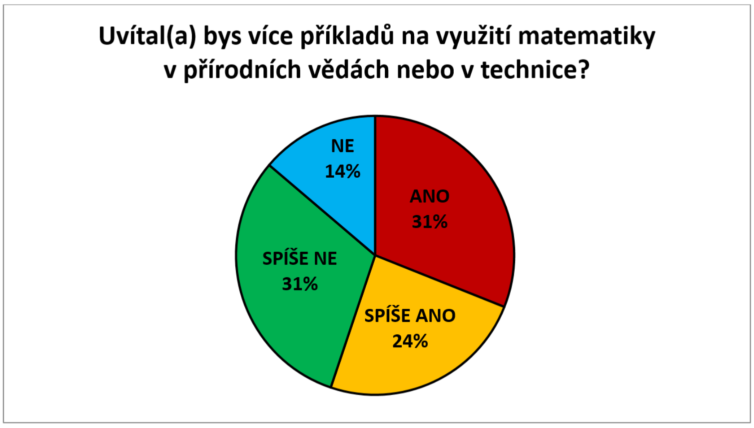
už ale byly odpovědi různorodější, neboť 7 z 13 odpovědělo, že je matematika spíše baví. Matematika baví 7 z 9 „jedničkářů“, jednoho spíše baví a jednoho spíše nebaví.



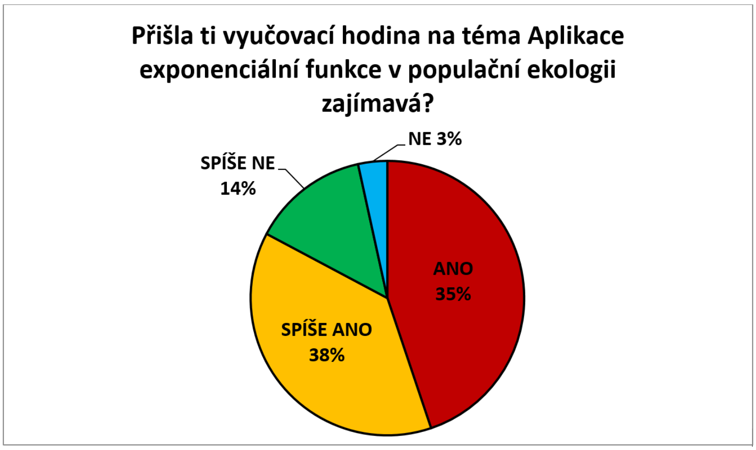
Následující otázka zjišťovala, zda studenti ve výuce počítají příklady na aplikaci matematiky. Výsledek byl velmi uspokojivý – 76 % uvedlo, že ano nebo spíše ano. Jen 3 % (tj. jeden student) se domnívá, že ne.



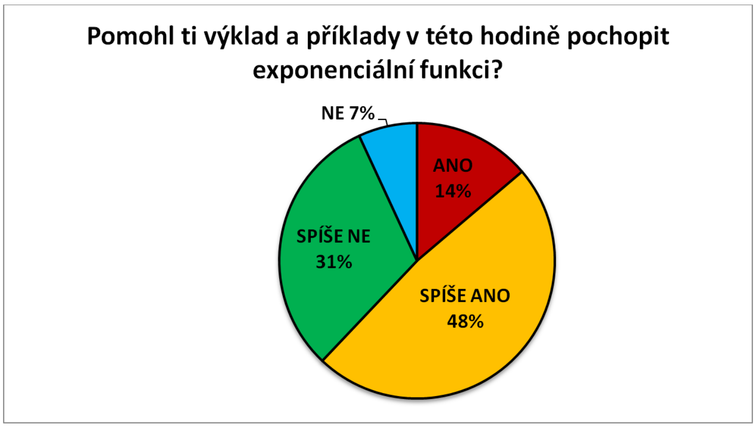
I tak se ale ukázalo, že 55 % dotazovaných by uvítalo nebo spíše uvítalo více příkladů na využití matematiky v přírodních vědách a technice.



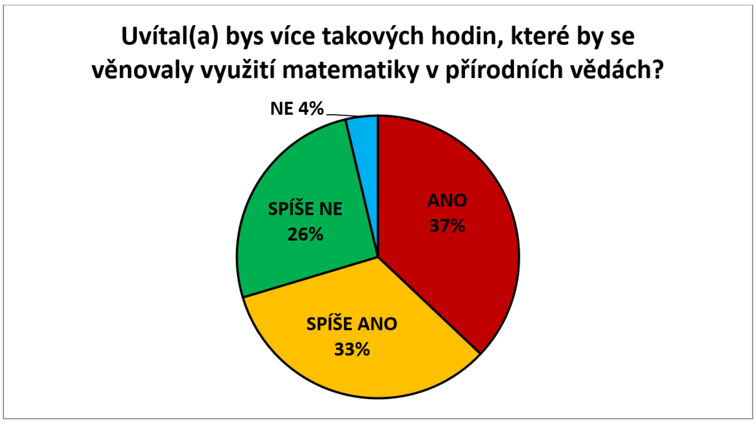
Druhá část dotazníku se týkala názoru studentů na mnou odučenou hodinu. Celkem 35 % dotazovaných shledalo moji hodinu jako zajímavou a 38 % jako spíše zajímavou, pouze 3 % (což odpovídá jednomu studentovi) přišla hodina nezajímavá. Hodnocení hodiny mne velmi potěšilo.



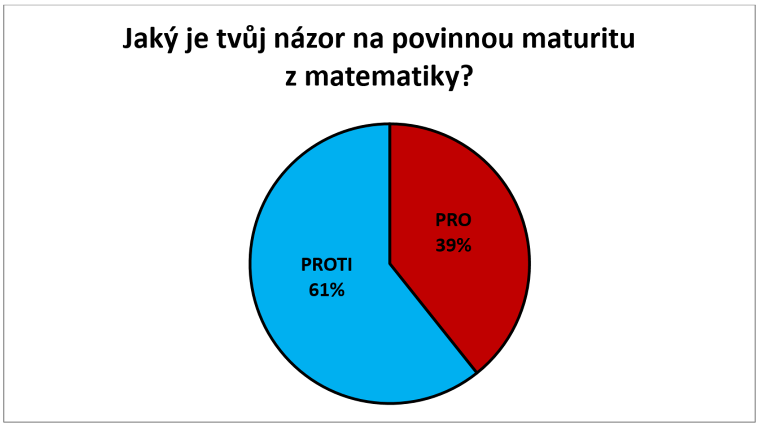
Celkem 62 % dotazovaných můj výklad a příklady pomohly nebo spíše pomohly pochopit již probíranou exponenciální funkci (pokud ovšem dotazovaní odpovídali upřímně). Mezi těmito 62 % byli i 3 ze 4 „čtyřkařů“, což mne mile překvapilo.



Předposlední otázka zjišťovala, zda by studenti uvítali více takovýchto hodin, které by je seznámily s využitím matematiky na poli přírodních věd. Odpovědi byly v 70 % kladné, tj. ano nebo spíše ano, a pouze jeden student by více takových hodin neuvítal.



Poslední otázka byla stejně jako u předešlého dotazníku otevřená a týkala se názoru studentů na povinnou státní maturitu z matematiky. Nebylo snadné jejich odpovědi striktně rozřadit do dvou skupin: „PRO“ a „PROTI“, neboť jen málokdo byl vyhraněně pro, resp. proti povinné maturitě. Většina odpovědí byla buď pro maturitu, ale ne pro stejnou úroveň na všech typech škol, nebo proti maturitě v současné podobě. Jeden z dotazovaných na otázku neodpověděl vůbec. Uvedené argumenty byly dosti podobné jako u dotazníku ke kapitole Aplikace podmíněné pravděpodobnosti v lékařské diagnostice.



Následující tabulka ukazuje četnosti jednotlivých odpovědí na dané otázky z obou částí dotazníku:

Obrázek 3.6: Četnosti jednotlivých odpovědí na otázky z obou částí dotazníku.

Četnosti jednotlivých odpovědí								
odpovědi	**část 1**				**část 2**			
	otázka 1	otázka 2	otázka 3	otázka 4	otázka 1	otázka 2	otázka 3	otázka 4
ANO	8	X	4	9	13	4	10	X
SPÍŠE ANO	11	X	18	7	11	14	9	X
SPÍŠE NE	9	X	6	9	4	9	7	X
NE	1	X	1	4	1	2	1	X
známka 1	X	9	X	X	X	X	X	X
známka 2	X	3	X	X	X	X	X	X
známka 3	X	13	X	X	X	X	X	X
známka 4	X	4	X	X	X	X	X	X
známka 5	X	0	X	X	X	X	X	X
PRO	X	X	X	X	X	X	X	11
PROTI	X	X	X	X	X	X	X	17
celkem	29	29	29	29	29	29	27	28

V poslední tabulce je uveden přehled všech dotazníků (označených číslem) a odpovědí na jednotlivé otázky. Lze si tedy například všimnout vztahu mezi známkou a oblibou matematiky (otázky 1 a 2 v 1. části dotazníku).

Obrázek 3.7: Přehled odpovědí u jednotlivých dotazníků.

číslo dotazníku	část 1				část 2			
	otázka 1	otázka 2	otázka 3	otázka 4	otázka 1	otázka 2	otázka 3	otázka 4
1	ANO	1	SPÍŠE ANO	SPÍŠE ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	PROTI
2	SPÍŠE NE	4	NE	NE	SPÍŠE ANO	ANO	SPÍŠE NE	PROTI
3	ANO	1	SPÍŠE NE	NE	SPÍŠE NE	NE	NE	PROTI
4	SPÍŠE ANO	3	SPÍŠE NE	SPÍŠE ANO	NE	SPÍŠE NE	SPÍŠE ENO	PROTI
5	SPÍŠE ANO	2	SPÍŠE NE	ANO	SPÍŠE ANO	SPÍŠE ANO	ANO	PRO
6	SPÍŠE NE	3	SPÍŠE NE	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	SPÍŠE NE	PROTI
7	SPÍŠE ANO	3	SPÍŠE ANO	ANO	SPÍŠE ANO	ANO	ANO	PRO
8	ANO	1	SPÍŠE ANO	SPÍŠE ANO	ANO	SPÍŠE ANO	ANO	PRO
9	ANO	1	SPÍŠE ANO	ANO	ANO	SPÍŠE NE	ANO	PROTI
10	SPÍŠE NE	4	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	PROTI
11	SPÍŠE ANO	2	SPÍŠE ANO	SPÍŠE NE	ANO	ANO	SPÍŠE NE	PRO
12	SPÍŠE NE	3	SPÍŠE ANO	SPÍŠE NE	SPÍŠE NE	SPÍŠE NE	SPÍŠE NE	PROTI
13	ANO	1	SPÍŠE ANO	ANO	ANO	ANO	ANO	PROTI
14	ANO	1	ANO	ANO	ANO	SPÍŠE NE	SPÍŠE ANO	PRO
15	SPÍŠE ANO	3	SPÍŠE ANO	SPÍŠE NE	ANO	SPÍŠE ANO	ANO	PRO
16	ANO	3	SPÍŠE ANO	SPÍŠE ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	PROTI
17	NE	4	SPÍŠE ANO	SPÍŠE NE	SPÍŠE NE	SPÍŠE NE	SPÍŠE NE	PROTI
18	SPÍŠE ANO	3	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE ANO	PÍŠE NE	PRO
19	SPÍŠE ANO	3	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	PROTI
20	SPÍŠE ANO	3	SPÍŠE ANO	SPÍŠE NE	SPÍŠE ANO	SPÍŠE NE	SPÍŠE NE	PROTI
21	SPÍŠE NE	3	SPÍŠE ANO	ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	PRO
22	SPÍŠE NE	1	ANO	ANO	ANO	SPÍŠE ANO	SPÍŠE ANO	PROTI
23	SPÍŠE ANO	3	SPÍŠE ANO	NE	SPÍŠE ANO	NE	SPÍŠE ANO	PROTI
24	ANO	1	ANO	NE	SPÍŠE ANO	SPÍŠE NE	SPÍŠE NE	PROTI
25	SPÍŠE ANO	1	SPÍŠE NE	SPÍŠE ANO	SPÍŠE ANO	SPÍŠE NE	ANO	PRO
26	SPÍŠE NE	4	SPÍŠE NE	ANO	SPÍŠE ANO	SPÍŠE ANO	ANO	-
27	SPÍŠE NE	3	SPÍŠE ANO	SPÍŠE NE	ANO	SPÍŠE ANO	SPÍŠE ANO	PRO
28	SPÍŠE NE	3	ANO	SPÍŠE NE	ANO	SPÍŠE ANO	ANO	PRO
29	SPÍŠE ANO	2	SPÍŠE ANO	ANO	ANO	SPÍŠE ANO	ANO	PROTI

4. Aplikace logických funkcí při matematickém modelování neuronu

Klíčová slova: Matematický model neuronu, logický neuron, model umělé nervové sítě, biologický neuron, logická funkce.

Předpoklad: Znalost základů výrokové logiky (viz např.: POLÁK, Josef. *Přehled středoškolské matematiky*. 10. vydání. Praha: Prometheus, 2015. ISBN 978-80-7196-458-2, s. 9 – 17).

V této kapitole se dozvíte:

- Jaká je stavba a funkce biologického neuronu.
- Co je to matematický model nervové sítě a umělý neuron, kde všude se tyto modely využívají.
- Co je to logická funkce a jak je možné ji graficky reprezentovat.
- Co je to McCullochův-Pittsův logický neuron, jak funguje a jak se graficky znázorňuje.
- Jak se matematický model neuronu liší od biologického neuronu a jaké jsou jeho výhody.

Použité značení:

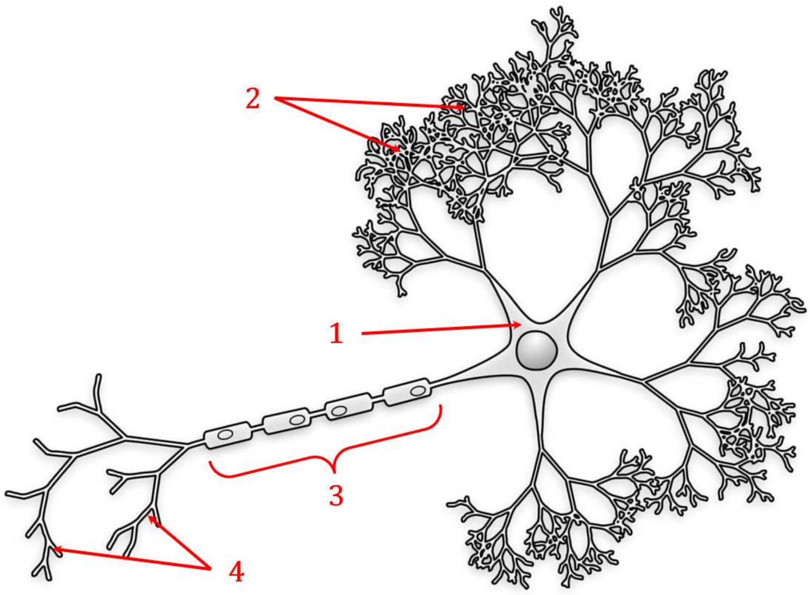
$E = \{x, y\}$	množina E, jejíž prvky jsou x, y
$x \in E$	x je prvkem množiny E
$E^2 = E \times E$	kartézská mocnina množiny E
$f(x_1, x_2, ..., x_n)$	funkce f proměnných $x_1, x_2, ..., x_n$
$\overline{x}$	negace prvku x
$\oplus$	operace exkluzivní součet
$\wedge$	operace konjunkce
$\vee$	operace disjunkce
$\Rightarrow$	operace implikace
$\Leftrightarrow$	operace ekvivalence
$\|X\|$	mohutnost (počet prvků) množiny X
$X \to Y$	zobrazení z množiny X do množiny Y

ϑ	prahový koeficient
ϵ	vnitřní potenciál

4.1 Motivace

Biologický neuron

Nervová soustava řídí a koordinuje činnost všech částí živého organismu, zároveň zprostředkovává propojení mezi jeho vnějším a vnitřním prostředím. Změny vnitřního nebo vnějšího prostředí ve formě fyzikálních nebo chemických sil se nazývají podněty. Podnět podráždí určitý receptor (součást smyslového orgánu) a dá tak vznik nervovému vzruchu, který se šíří dostředivými vlákny nervů do centrální nervové soustavy (míchy a mozku), kde dochází k jejich shromažďování, zpracování a vyhodnocování. Následně vzruch pokračuje po odstředivých vláknech směrem od centrální nervové soustavy k výkonnému orgánu (svalu nebo žláze), který provede reakci organismu na podráždění (např. zúžení zornice po podráždění oka světlem).[14]
Základní stavební a funkční jednotkou nervové soustavy je **nervová buňka (neuron)**. V těle se nachází kolem 100 miliard (10^{11}) vysoce specializovaných neuronů, v mozku je jich přibližně desetina. Vzájemným propojením neuronů vznikají nervové provazce, v centrální nervové soustavě dokonce neurony vytvářejí komplikovanou prostorovou **nervovou síť**.[15]



Obrázek 4.1: Schéma nervové buňky (převzato z [16]): 1 – buněčné tělo s jádrem, 2 – dendrity, 3 – axon a 4 – nervová zakončení.

Buněčné tělo neuronu představuje jakési ústředí, do kterého jsou vedeny impulzy po dostředivých vláknech z ostatních neuronů nebo smyslových buněk. Tato krátká početná vlákna nazýváme dendrity. Dlouhé odstředivé vlákno axon (neurit) poté vede vzruch směrem od těla neuronu. Konec axonu se mnohonásobně větví a dotýká se dalších nervových buněk nebo výkonných orgánů. Neuron má zpravidla pouze jeden axon a může dosahovat délky až jednoho metru. Neurony si mezi sebou předávají vzruchy chemickou cestou pomocí tzv. neurotransmiterů (přenašečů). Po neuronech se vzruchy šíří v podobě elektrického akčního potenciálu, jehož zdrojem jsou rozdílné koncentrace iontů (resp. jimi nesené rozdílné náboje) vně a uvnitř buněčné membrány. Rychlost šíření vzruchů je v nervové soustavě rozdílná a může dosahovat až rychlosti přes $100\ \mathrm{m} \cdot \mathrm{s}^{-1}$.[15]

Model neuronové sítě a umělý neuron

Matematický model neuronové sítě je výpočetní model využívaný v umělé inteligenci[1], vytvořený na základě chování biologické neuronové sítě. Tyto sítě slouží k distribuovanému paralelnímu zpracování dat. Umělá neuronová síť vzniká seskupením a vzájemným propojením **umělých (formálních) neuronů** majících libovolný počet vstupů a jediný výstup; jejich vzorem v chování je biologický neuron. Formální neurony si mezi sebou předávají signály a transformují je tzv. přenosovými funkcemi.[17]

Neuronová síť dokáže v aktivním režimu realizovat požadovanou funkci. V určitém smyslu se tedy jedná o univerzálním výpočetní prostředek jako např. klasický počítač. Vše, co spočítá klasický počítač, dokáže spočítat také neuronová síť. Její výhodou a zároveň odlišností oproti klasickému počítači je **schopnost učit se a zobecňovat (generalizovat)**, což je zároveň typickou vlastností lidské inteligence.[17]

Modely neuronových sítí nachází uplatnění všude tam, kde nelze použít klasické počítače – nejčastěji se jedná o praktické problémy, kdy nám není znám jejich algoritmus nebo je jejich analytický popis pro počítač příliš složitý. To však neznamená, že by tyto modely mohly počítače zcela nahradit, neboť při mechanických a jednoduše analyticky popsatelných výpočtech počítače co se týče rychlosti a přesnosti bezkonkurenčně vítězí.[18]

[1]Umělá inteligence je obor informatiky zabývající se tvorbou strojů vykazujících známky inteligentního chování.

Neuronové sítě mají mnoho aplikací v různých oblastech, výčtem uveďme ty nejdůležitější:

- **rozpoznávání obrazců** (např. při rozpoznávání naskenovaného psaného nebo tištěného textu)
- **transformace signálů** (např. systém NETtalk sloužící pro převod anglicky psaného textu na mluvený signál)
- **predikce a případně následné rozhodování** (např. meteorologické předpovědi, vývoj cen akcií na burze, spotřeba elektrické energie)
- **řízení složitých zařízení v dynamicky se měnících podmínkách** (neuronové sítě např. umožňují autopilotu se učit kopírováním techniky pilotáže letce nebo řidiče)
- **komprese dat** (např. pro přenos televizního signálu, telekomunikaci)
- **analýza signálů** (např. EKG, EEG)[18]
- **prohlubování znalostí o fungování nervových soustav živých organismů**
- **expertní systémy** (systém EXPSYS v energetice a lékařství, např. v lékařské aplikaci jsou vstupem do neuronové sítě zakódované příznaky určitých onemocnění a výsledky vyšetření, jejím výstupem je doporučená léčba)[17]

4.2 Logické funkce a jejich grafické znázornění

Dříve než se začneme zabývat vlastním logickým neuronem a matematickým modelem neuronu, je potřeba znát několik základních poznatků z matematické logiky – zejména logické funkce, jejich vlastnosti a grafické reprezentace.

Nejprve si uveďme definici dvouhodnotové logické funkce:

Definice 4.2.1. ([19], s. 35) **[Logická funkce.]** Nechť $E^2 = \{0, 1\}$. Pak funkce $f(x_1, x_2, ..., x_n)$ definovaná zobrazením $f : E^2 \times E^2 \times ... \times E^2 \to E^2$ se nazývá logická funkce n proměnných (nebo také Booleova funkce n proměnných). Množinu všech takových funkcí budeme označovat P_2.

Proměnné $x_1, x_2, ..., x_n$ funkce f nabývající pouze hodnot 0 a 1 budeme označovat jako **logické proměnné**. Každá logická funkce s n proměnnými přiřadí n-tici nul a jedniček hodnotu nula anebo jedna.

Definice 4.2.2. ([20], s. 30) [**Pravdivostní ohodnocení.**] Pravdivostním ohodnocením logických proměnných rozumíme jakékoliv přiřazení hodnoty 0 („nepravda“) anebo 1 („pravda“) každé z těchto logických proměnných.

Věta 4.2.1. ([19], s. 18) Nechť X a Y jsou dvě konečné množiny takové, že $|X| = k$ a $|Y| = n$ (tedy množina X je k-prvková a množina Y n-prvková). Pak počet všech možných zobrazení $f : X \to Y$ je roven n^k.

Věta 4.2.2. ([19], s. 35) Počet všech logických funkcí z množiny P_2, závisejících na n proměnných, je roven 2^{2^n}.

Důkaz: Máme-li logickou funkci s n proměnnými $f(x_1, x_2, ..., x_n)$, pak je počet všech různých $n-$tic proměnných, resp. počet všech různých hodnot funkce, roven počtu všech variací n-té třídy ze dvou prvků s opakováním. Z kombinatoriky víme, že pro počet všech těchto variací platí $V_n'(2) = 2^n$. Logická funkce f je dle Definice 4.2.1 zobrazením množiny, která má 2^n prvků do množiny obsahující právě 2 prvky (pro každou n-tici proměnných jsou 2 možné funkční hodnoty). Počet všech takovýchto zobrazení je podle Věty 4.2.1 roven 2^{2^n}. $\square$

Logické funkce lze zadávat buď tabulkou **pravdivostních hodnot** nebo **analyticky**. Tabulka pravdivostních hodnot je praktický a přehledný způsob, jak snadno určit všechny možné hodnoty logické funkce.

Tabulka 4.1: Tabulka pravdivostních hodnot logické funkce $f(x_1, x_2, ..., x_n)$ s n proměnnými.

x_1	...	x_{n-1}	x_n	$f(x_1, ..., x_{n-1}, x_n)$
0	...	0	0	$f(0, ..., 0, 0)$
0	...	0	1	$f(0, ..., 0, 1)$
0	...	1	1	$f(0, ..., 1, 1)$
...	...	...	...	...
1	...	1	1	$f(1, ..., 1, 1)$

Prvních n sloupců tabulky přísluší jednomu ohodnocení $n$tice logických proměnných $x_1, x_2, ..., x_n$. V posledním sloupci tabulky se nachází pravdivostní hodnota příslušné funkce $f(x_1, ..., x_{n-1}, x_n)$ při tomto ohodnocení $n-$tice proměnných.

Obecně tabulka pravdivostních hodnot logické funkce s n proměnnými má 2^n řádků a alespoň $n + 1$ sloupců (pokud zápis nějak nezjednodušíme). [20]

Příklad 4.2.1. Jako příklad si uveďme logickou funkci dvou proměnných x, y s názvem exkluzivní součet, nebo také nonekvivalence. Její analytický předpis je tvaru:

$$f(x, y) = x \oplus y = (x \wedge \overline{y}) \vee (\overline{x} \wedge y), \;\; x, y \in \{0, 1\}.$$

Tabulkou pravdivostních hodnot bychom tuto funkci zadali následovně:

x	y	$f(x, y)$
1	1	0
1	0	1
0	1	1
0	0	0

Některými konkrétními typy logických funkcí se budeme zabývat později.

Určitě jste si všimli v předchozích příkladech, že logické funkce jsou tvořeny logickými proměnnými, které jsou spojeny **logickými operacemi**, dobře známými ze základů matematické logiky. V předchozích příkladech se konkrétně vyskytovaly operace negace, konjunkce a disjunkce. Tyto operace a také operaci implikace budeme nadále potřebovat.
Logické operace mají určitou prioritu, která udává, v jakém pořadí jsou jednotlivé logické operace prováděny. Vyhodnocování výrazu se provádí klasicky zleva doprava. Pravidlo priority operací je následující – negace, konjunkce, disjunkce. Pořadí vyhodnocování operací lze změnit prostřednictvím závorek. Pravdivost těchto operací vyplyne z pravdivosti logických funkcí, které si nyní představíme.

Logická funkce byla v úvodu podkapitoly definována obecně pro n proměnných (viz Definice 4.2.1). Pro naši potřebu nyní postačí ukázat negaci – logickou funkci jedné proměnné a tyto funkce dvou proměnných – konjunkci, disjunkci a implikaci:

1. **negace**, je taková funkce, jejíž pravdivostní hodnota se rovná opačné pravdivostní hodnotě vstupní proměnné a jejíž předpis je $f_1(x) = \overline{x}$, resp. $f_1(x) = \neg x$,

2. **konjunkce** (**logický součin**, „a“) je taková funkce, která nabývá pravdivostní hodnoty 1, právě když obě vstupní pravdivostní hodnoty proměnných mají současně hodnotu 1, předpis konjunkce má tvar $f_2(x, y) = x \wedge y$,

3. **disjunkce** (**logický součet**, „nebo“) je taková funkce, která nabývá pravdivostní hodnoty 1, právě když alespoň jedna ze vstupních pravdivostních hodnot proměnných nabývá hodnoty 1, její předpis je $f_3(x, y) = x \vee y$,

4. **implikace** („jestliže..., pak“) je taková funkce, která nabývá pravdivostní hodnoty 1 ve všech případech kromě případu, kdy vstupní pravdivostní hodnota první proměnné je 1 a druhé 0, její předpis je $f_4(x, y) = x \Rightarrow y$,

kde $x, y \in \{0, 1\}$. [19]

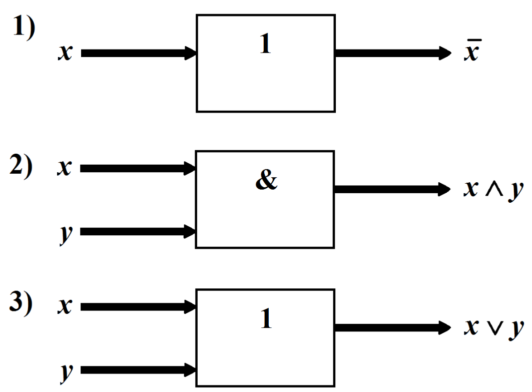
Snadno lze určit všechny pravdivostní hodnoty těchto funkcí použitím tabulky pravdivostních hodnot:

Tabulka 4.2: Tabulka pravdivostních hodnot negace, konjunkce, disjunkce a implikace.

x	y	$f_1 = \overline{x}$	$f_2 = x \wedge y$	$f_3 = x \vee y$	$f_4 = x \Rightarrow y$
1	1	0	1	1	1
1	0	0	0	1	0
0	1	1	0	1	1
0	0	1	0	0	1

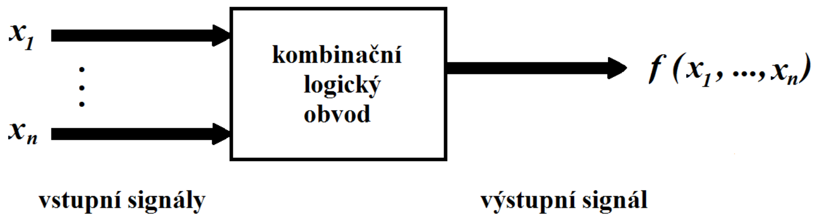
Logické funkce lze také znázorňovat **graficky**. Grafické reprezentace logických funkcí se nazývají **logické členy** nebo také **logická hradla**.

Logická hradla pro funkce negace, konjunkce a disjunkce vypadají následovně:



Obrázek 4.2: Logická hradla pro 1) negaci, 2) konjunkci a 3) disjunkci. (Podle [21], s. 220).

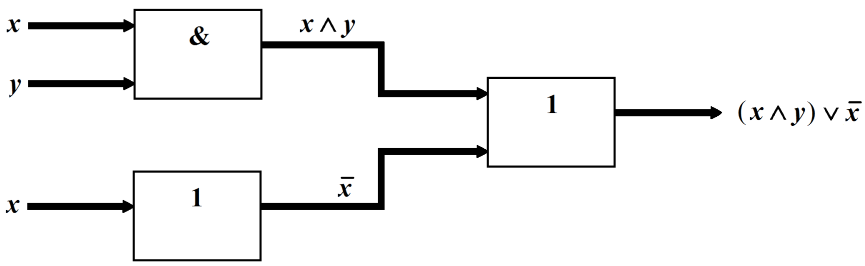
Logická hradla tvoří stavební prvky složitějších stavebních systémů – tzv. **kombinačních logických obvodů** (viz schéma 4.3).



Obrázek 4.3: Schéma kombinačního logického obvodu.

Kombinační logické obvody jsou speciálním případem obecnějších logických obvodů. Představují fyzikální realizace logických funkcí, které přetváří vstupní signály v jeden výstupní signál podle žádané činnosti daného zařízení. Všechny tyto signály nabývají pouze hodnot 0 (signál neprochází zařízením) a 1 (signál prochází zařízením). Výstupní signály kombinačních obvodů jsou jednoznačně určeny kombinací současně působících vstupních signálů. Každý takovýto kombinační obvod má n vstupů, právě jeden výstup a představuje logickou funkci $f : 2^n \to 2$. [21]

Příklad 4.2.2. Pokud bychom chtěli zakreslit logickou funkci $f(x, y) = (x \wedge y) \vee \overline{x}$ pomocí hradel, kombinační logický obvod reprezentující tuto funkci by mohl vypadat např. takto:



Obrázek 4.4: Kombinační obvod pro funkci $f(x, y) = (x \wedge y) \vee \overline{x}$.

Poznámka 4.2.1. Z logických funkcí můžeme jejich spojováním vytvářet **logické formule**. Příkladem logické formule je např. funkce f v předchozím příkladě, která vznikla spojením funkce konjunkce $f_1 = x \wedge y$ a negace $f_2 = \overline{x}$ pomocí logické operace disjunkce.

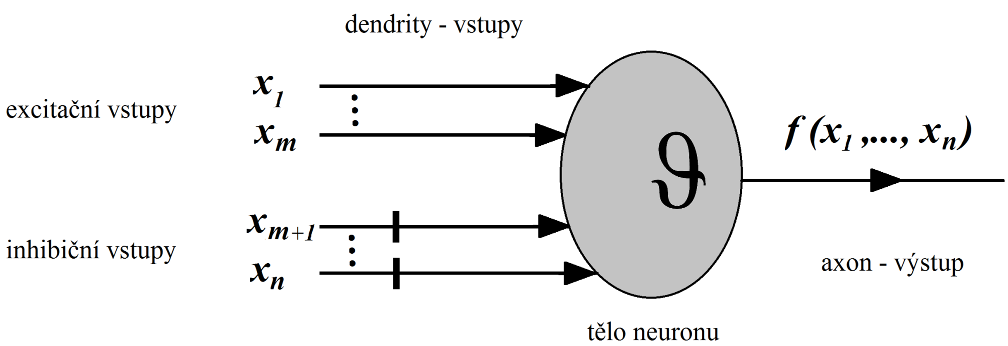
4.3 Matematický model neuronu

Úvodem pár slov k historii neurovýpočtů:

Základy oboru neuronových sítí položili v roce 1943 Warren McCulloch a Walter Pitts svou prací *„A logical calculus of the ideas immanent to nervous activity"*, ve které prezentovali velmi jednoduchý **matematický model neuronu**. McCulloch a Pitts uvedli, že nejjednodušší neuronové sítě mohou spočítat jakoukoliv aritmetickou nebo logickou funkci. Jejich práce v budoucnosti inspirovala mnohé badatele.[18]

Logický neuron (McCullochův-Pittsův logický neuron), základní jednotka McCullochových a Pittsových neuronových sítí, vlastně představuje výpočetní jednotku, která má dva stavy (tj. je binární) – stav 1 (ozn. „aktivní") nebo 0 (ozn. „neaktivní"), a proto ji lze vyložit jako relé[2].[17]

Logický neuron, jehož struktura je schematicky znázorněná na obr. 4.5, je možné si v jeho nejjednodušší podobě představit jako zařízení s několika vstupy (modelujícími dendrity) a jedním výstupem (modelujícím axon), který se však může dále rozvětvovat a být následně zapojen do většího či menšího počtu dalších neuronů. Vstupy se schematicky zakreslují jako spojnice s šipkami směřujícími k symbolu těla neuronu a výstup se zakresluje pomocí spojnic vycházejících z těla neuronu. [22]



Obrázek 4.5: Nejběžnější schéma McCullochova a Pittsova neuronu s prahem ϑ. (Upraveno podle [17], s. 18.)

[2]jednoduché elektrické zařízení sloužící ke spínání signálu

Dendritický systém logického neuronu zahrnuje vstupy dvojího druhu:

1. **excitační** – zesilující odezvu (na obr. 4.5 reprezentovány binárními proměnnými $x_1, ...,$ x_m),
2. **inhibiční** – zeslabující (na obr. 4.5 reprezentovány proměnnými $x_{m+1}, ..., x_n$ a označovány kolmou čárkou přes spojnici).[17]

Tělo neuronu je charakterizováno **prahovým koeficientem** neboli prahem (ozn. ϑ), který je vždy větší nebo roven nule. Jedná se o minimální počet aktivních excitačních vstupů (vstupů s jednotkovou aktivitou), které uvedou v aktivní stav celý logický neuron – na výstupu neuronu se tedy objeví signál.[22]

Vnitřní potenciál (ozn. ξ) neuronu je definován jako rozdíl mezi sumou excitačních vstupů a sumou inhibičních vstupů:[17]

$$\xi = \sum_{i=1}^{m} x_i - \sum_{j=m+1}^{n} x_j \tag{4.1}$$

Hodnota vnitřního potenciálu ξ po dosažení prahové hodnoty ϑ indukuje **výstup** (stav aktivity) logického neuronu (ozn. $f(x_1, ..., x_n)$) modelující elektrický impuls axonu.[18]

Platí tedy, že neuron je

- **aktivní**, tzn. $f(x_1, ..., x_n) = 1$, jestliže $\xi \geq \vartheta$,
- **neaktivní**, tzn. $f(x_1, ..., x_n) = 0$, jestliže $\xi < \vartheta$.

Zjištění výstupu neuronu pomocí předchozích zjištění demonstruje následující příklad:

Příklad 4.3.1. Mějme neuron s třemi excitačními vstupy x_1, x_2, x_3 a dvěma inhibičními vstupy x_4, x_5. Jeho prahový koeficient ϑ je 1. Rozhodněte, zda-li bude neuron aktivní, když

1. $x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 1, x_5 = 0$?
2. $x_1 = 0, x_2 = 1, x_3 = 0, x_4 = 0, x_5 = 1$?

Řešení:

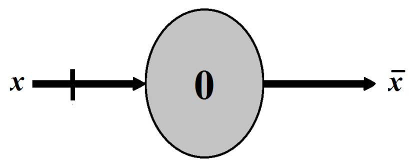
1. Ano, bude aktivní, neboť dosazením do vzorce (4.1) získáme: $\xi = 1$ a tedy $\xi \geq \vartheta$.
2. Ne, nebude aktivní, neboť dosazením do vzorce (4.1) získáme: $\xi = 0$ a tedy $\xi < \vartheta$.

McCulloch a Pitts ve své práci poukázali na fakt, že logické neurony tak, jak je zavedli, mohou jednoznačně představovat některé logické funkce.[3] Proto se dříve předpokládalo, že na stejném principu funguje i lidský mozek. Nyní již při současné úrovni poznání víme, že by takovéto chování při obrovském počtu mozkových neuronů (10^{10}) a spojů mezi nimi (minimálně $10^4 \cdot 10^{10}$) bylo nereálné a mozek byl příliš těžkopádný.[22]

Logický neuron je tedy zjednodušeným idealizovaným modelem skutečného biologického neuronu, za což může být právem kritizován odborníky z řad biologů a neurofyziologů. Zachovává si však podstatné charakteristické vlastnosti biologického neuronu a zanedbává ty nepodstatné. Právě jednoduchost logického neuronu je jeho obrovskou výhodou, neboť umožňuje analyzovat nervovou síť, což by bylo v případě zahrnutí množství fyziologických hledisek velmi složité. Často je zcela nezbytné si reálný svět zjednodušovat, abychom jej dokázali modelovat a podrobovat analýze.[22]

Vraťme se ale k myšlence **propojení logického neuronu a logických funkcí**, které se dodnes využívá na poli matematického modelování. Obecné logické neurony pro implementaci logických funkcí negace, disjunkce, konjunkce a implikace vypadají následovně:[22]

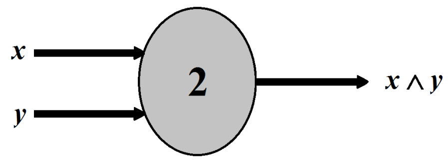
- **Negaci** představuje logický neuron s jedním inhibičním vstupem, jedním výstupem a prahovou hodnotou $\vartheta = 0$. Neuron je aktivní, pokud je jeho inhibiční vstup neaktivní.



Obrázek 4.6: Logický neuron reprezentující negaci.

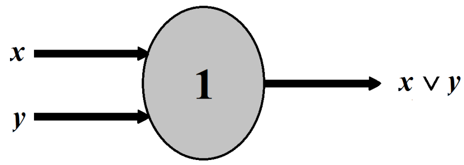
- **Konjunkci** dvou proměnných představuje logický neuron s dvěma excitačními vstupy, jedním výstupem a prahovou hodnotou $\vartheta = 2$. Tento neuron je aktivní, pokud jsou oba jeho excitační vstupy aktivní.

[3] Jedná se o tzv. lineárně separabilní funkce – např. negace, konjunkce, disjunkce a implikace. Lineárně separabilní není např. funkce exkluzivní součet uvedená v Příkladu 4.2.1.



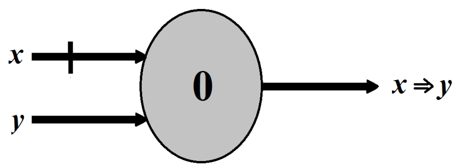
Obrázek 4.7: Logický neuron reprezentující konjunkci.

- **Disjunkci** dvou proměnných představuje logický neuron s dvěma excitačními vstupy, jedním výstupem a prahovou hodnotou $\vartheta = 1$. Tento neuron je aktivní, pokud je aspoň jeden z jeho excitačních vstupů aktivní.



Obrázek 4.8: Logický neuron reprezentující disjunkci.

- **Implikaci** dvou proměnných představuje logický neuron s jedním excitačním vstupem, jedním inhibičním vstupem a prahovou hodnotou $\vartheta = 0$. Neuron je aktivní ve všech případech kromě případu, kdy je jeho inhibiční vstup aktivní a excitační vstup neaktivní. Výstup tohoto neuronu lze charakterizovat také funkcí $f = \overline{x} \vee y$, kde $x, y \in \{0, 1\}$, která je ekvivalentní s funkcí implikace.

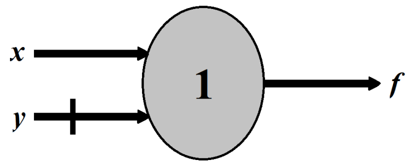


Obrázek 4.9: Logický neuron reprezentující implikaci.

Prahovou hodnotu lze tedy matematicky vyjádřit jako minimální počet proměnných s pravdivostní hodnotou 1, který je potřebný pro to, aby funkční hodnota logické funkce byla 1.

Předchozí zjištění si názorně ukážeme na řešených příkladech:

Příklad 4.3.2. Pro logický neuron, kde $x, y \in \{0, 1\}$, určete:



1. Kolik má excitačních a inhibičních vstupů?
2. Jaká je jeho prahová hodnota?
3. Za jakých podmínek je tento neuron aktivní?
4. Jakou logickou funkci dvou proměnných neuron reprezentuje?

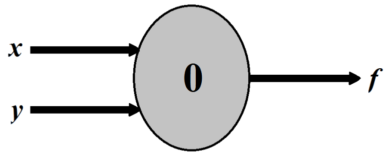
Řešení:

1. Tento neuron má jeden excitační a jeden inhibiční vstup, což je zřejmé ze zadání.
2. Prahovou hodnotu $\vartheta = 1$ lze také ihned vyčíst ze zadání.
3. Rozhodujeme-li o aktivitě neuronu, můžeme využít porovnání vnitřního potenciálu a prahové hodnoty. Navíc z hodnoty prahového koeficientu víme, že potřebujeme aktivní excitační vstup. Neuron bude aktivní, když $\xi \geq \vartheta$, což se stane pouze v případě, pokud bude aktivní jeho excitační vstup a neaktivní inhibiční vstup, viz názorná tabulka:

x	y	ξ	$f(x, y)$	stav
1	1	0	0	neaktivní
1	0	1	1	**aktivní**
0	1	-1	0	neaktivní
0	0	0	0	neaktivní

4. Z předchozí tabulky je patrné, že funkce bude mít na výstupu pravdivostní hodnotu 1 pouze pokud $x = 1$ a $y = 0$, a tedy že neuron reprezentuje logickou funkci $f(x, y) = x \wedge \overline{y}$, kde $x, y \in \{0, 1\}$.

Příklad 4.3.3. Pro logický neuron, kde $x, y \in \{0, 1\}$, určete:



1. Kolik má excitačních a inhibičních vstupů? Jaká je jeho prahová hodnota?
2. Za jakých podmínek je tento neuron aktivní?
3. Jakou logickou funkci dvou proměnných neuron reprezentuje?

Řešení:

1. Tento neuron má dva excitační vstupy a nemá žádný inhibiční vstup. Prahová hodnota tohoto neuronu je $\vartheta = 0$.
2. V tomto případě lze na základě prahové hodnoty snadno rovnou rozhodnout, že neuron bude aktivní ve všech případech.[4] Výpočtem a pomocí tabulky ověříme, že platí $\xi \geq \vartheta$ pro všechna x, y:

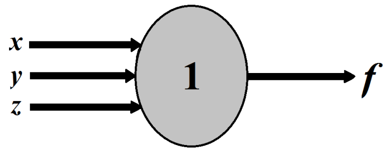
x	y	ξ	$f(x,y)$	stav
1	1	2	1	**aktivní**
1	0	1	1	**aktivní**
0	1	1	1	**aktivní**
0	0	0	1	**aktivní**

3. Logická funkce, kterou tento neuron reprezentuje lze vyjádřit jako $f(x, y) = 1$, $\forall x, y \in \{0, 1\}$.

Poznámka 4.3.1. Logickou funkci, která je vždy pravdivá, tj. $f(x, y) = 1$, $\forall x, y \in \{0, 1\}$ (viz Příklad 4.3.3), nazýváme **tautologie**. Naopak logickou funkci, která je vždy nepravdivá $f(x, y) = 0$, $\forall x, y \in \{0, 1\}$ nazýváme **kontradikce**.

[4] Neuron, který je aktivní i v případě, když není aktivní žádný z jeho vstupů, se nazývá spontánně aktivní.[22]

Příklad 4.3.4. Pro logický neuron, kde $x, y, z \in \{0, 1\}$, určete:



1. Za jakých podmínek je tento neuron aktivní?

2. Jakou logickou funkci tří proměnných neuron reprezentuje? Funkci zadejte tabulkou i analyticky.

Řešení:

1. Prahová hodnota neuronu je $\vartheta = 1$ a všechny jeho vstupy jsou excitační, což nám usnadňuje práci. Aniž bychom porovnávali ξ a ϑ, můžeme proto rovnou říci, že neuron aktivní, pokud bude aktivní alespoň jeden z jeho vstupů.

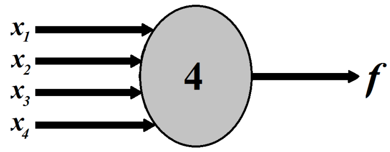
2. Vytvoříme si příslušnou tabulku:

x	y	z	$f(x, y, z)$
1	1	1	1
1	1	0	1
1	0	1	1
0	1	1	1
1	0	0	1
0	1	0	1
0	0	1	1
0	0	0	0

Funkce zadaná analyticky má tvar: $f(x, y, z) = x \vee y \vee z$, kde $x, y, z \in \{0, 1\}$.

Příklad 4.3.5. Zakreslete neuron, který má čtyři excitační vstupy $x_1, x_2, x_3, x_4 \in \{0, 1\}$ a jehož prahová hodnota je $\vartheta = 4$. Určete za jakých podmínek bude aktivní a jakou logickou funkci čtyř proměnných tento neuron reprezentuje.

Řešení:



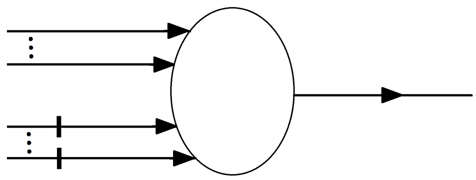
Neuron bude na základě své prahové hodnoty aktivní pouze v případě, pokud budou aktivní všechny jeho vstupy. Je zřejmé, že se jedná o funkci $f(x_1, x_2, x_3, x_4) = x_1 \wedge x_2 \wedge x_3 \wedge x_4$, kde $x_1, x_2, x_3, x_4 \in \{0, 1\}$.

4.4 Pracovní list

Téma: Aplikace logických funkcí při matematickém modelování neuronu

Jméno a třída: Datum:

Příklad č. 1 Doplň do schématu logického neuronu tyto jeho části: axon (výstup), dendrity (vstupy), excitační vstupy, inhibiční vstupy, tělo neuronu; a vyznač prahový koeficient ϑ.



Příklad č. 2 Škrtni nehodící se pojmy v textu o logickém neuronu:

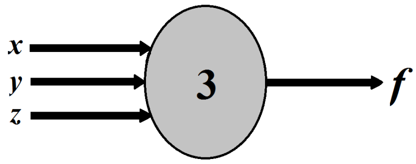
Logický neuron lze interpretovat jako jednoduché elektrické zařízení mající *několik vstupů/jeden vstup* a *několik výstupů/jeden výstup*. Práh neboli prahový koeficient je *minimální/maximální* počet aktivních vstupů potřebných k uvedení neuronu do *aktivního/neaktivního* stavu. Hlavní *nevýhodou/výhodou* modelu logického neuronu oproti biologickému neuronu je jeho jednoduchost, která *umožňuje/znemožňuje* jeho efektivní analýzu.

Příklad č. 3

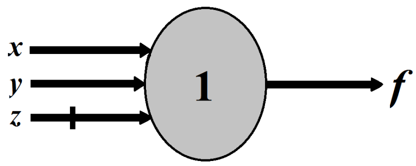
1. Zakresli logický neuron, který má jeden excitační vstup x, jeden inhibiční vstup y, kde $x, y \in \{0, 1\}$, a jehož prahová hodnota je $\vartheta = 0$.

2. Urči, za jakých podmínek bude tento neuron aktivní a jakou funkci na výstupu reprezentuje.

Příklad č. 4 Za jakých podmínek bude logický neuron na následujícím obrázku aktivní? Jaká logická funkce tří proměnných $x, y, z \in \{0, 1\}$ je na výstupu f reprezentována tímto neuronem, kde $\vartheta = 3$?

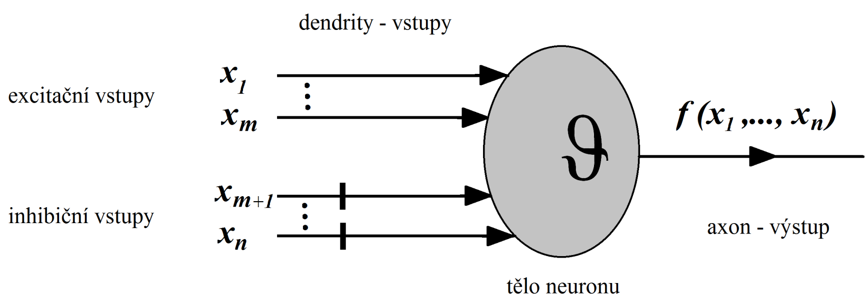


Příklad č. 5 Za jakých podmínek bude logický neuron na následujícím obrázku aktivní? Jakou logickou funkci tří proměnných $x, y, z \in \{0, 1\}$ tento neuron na výstupu f reprezentuje, když $\vartheta = 1$? Postačí, když funkci zadáš tabulkou.



Řešení pracovního listu:

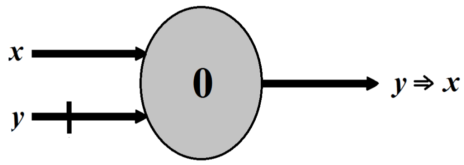
Příklad č. 1 Doplněné schéma logického neuronu vypadá následovně



Příklad č. 2 Škrtá se: jeden vstup, několik výstupů, maximální, neaktivního, nevýhodou, znemožňuje.

Příklad č. 3

1. Schéma daného neuronu je tvaru



2. Výstup f bude aktivní, ve všech případech kromě případu, kdy je jeho excitační vstup aktivní a inhibiční vstup neaktivní. Funkce na výstupu je tvaru $f = x \vee \overline{y} = y \Rightarrow x$, kde $x, y \in \{0, 1\}$:

x	y	$f(x,y)$
1	1	1
1	0	1
0	1	0
0	0	1

Příklad č. 4 Výstup f bude aktivní, když bude platit $\xi \geq \vartheta$, což nastane pouze v případě, pokud budou aktivní všechny jeho tři vstupy. Příslušná logická funkce je potom tvaru $f(x, y, z) = x \wedge y \wedge z$, kde $x, y, z \in \{0, 1\}$:

x	y	z	ξ	$f(x, y, z)$
1	1	1	3	1
1	1	0	2	0
1	0	1	2	0
0	1	1	2	0
1	0	0	1	0
0	1	0	1	0
0	0	1	1	0
0	0	0	0	0

Příklad č. 5 Výstup f bude aktivní, když bude platit $\xi \geq \vartheta$. Tato situace nastane, pokud budou aktivní buď oba excitační vstupy současně, nebo bude aktivní alespoň jeden z nich a současně neaktivní inhibiční vstup. Funkce má tvar $f(x, y, z) = (x \wedge y) \vee [(x \vee y) \wedge \overline{z}]$, kde $x, y, z \in \{0, 1\}$:

x	y	z	ξ	$f(x, y, z)$
1	1	1	1	1
1	1	0	2	1
1	0	1	0	0
0	1	1	0	0
1	0	0	1	1
0	1	0	1	1
0	0	1	-1	0
0	0	0	0	0

5. Aplikace binomické věty a binomického rozdělení v genetice

Klíčová slova: Binomická věta, binomické rozdělení, populační genetika, Hardy-Weinbergova rovnováha.

Předpoklad: Znalost základů kombinatoriky (viz např.: POLÁK, Josef. *Přehled středoškolské matematiky*. 10. vydání. Praha: Prometheus, 2015. ISBN 978-80-7196-458-2, s. 312 – 324) a znalost základních genetických pojmů – dědičnost, gen, alela, genotyp, fenotyp, štěpný poměr, Mendelovy zákony (viz např.: ŠMARDA, Jan. *Genetika pro gymnázia.* Praha: Fortuna, 2003. ISBN 80-7168-851-7, s. 20 – 33).

V této kapitole se dozvíte:

- Co je kombinační číslo a jaké jsou jeho vlastnosti. Co je binomická věta a binomické rozdělení.
- Jak lze snadno pomocí binomické věty a binomického rozdělení zjistit pravděpodobnost narození určitého počtu chlapců a dívek v jedné rodině.
- Co je to panmiktická populace, jaký je princip Hardy-Weinbergova zákona a jak se určuje frekvence alel a genotypů v populaci v Hardy-Weinbergově rovnováze.

Použité značení:

$\binom{n}{k}$	kombinační číslo, čteme „n nad k“
$n!$	faktoriál čísla n
$P(X)$	pravděpodobnost jevu X
A	dominantní alela A
a	recesivní alela a
p	frekvence dominantní alely A v populaci
q	frekvence recesivní alely a v populaci

5.1 Motivace

Ze středoškolské matematiky si každý jistě pamatuje známé vzorečky $(a \pm b)^2 = a^2 \pm 2ab + b^2$ a $(a \pm b)^3 = a^3 \pm 3a^2b + 3ab^2 \pm b^3$, které se odvozují z binomické věty. Aplikace binomické věty však nespočívá pouze ve zjištění určitého rozvoje $(a \pm b)^n$ nebo jeho $k-$tého členu, nýbrž nachází své využití v mnohých odvětvích, např. ve fyzice, v informatice, architektuře, ekonomii, meteorologii, ale také v genetice.

V této kapitole si ukážeme, jak za pomocí binomické věty, binomického rozvoje a Pascalova trojúhelníku snadno určíme pravděpodobnost narození k chlapců a m děvčat v rodině, která má $(k+m)$ dětí.

Dále si ukážeme, jak zjistíme frekvenci určité alely, genotypu nebo fenotypu v populacích, v nichž je zachována Hardy-Weinbergova rovnováha, např. jak vypočítat, kolik procent lidí v dané populaci je přenašečem mutantní alely způsobující určité onemocnění.

5.2 Binomická věta a binomické rozdělení

Definice 5.2.1 ([23], s. 27). [**Kombinační číslo.**] Kombinační číslo $\binom{n}{k}$, kde $k, n \in \mathbb{N}_0$ a $n \geq k$, je takové číslo, pro které platí

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{5.1}$$

Kombinační číslo $\binom{n}{k}$ udává počet kombinací, tedy počet způsobů, jak vybrat k-prvkovou podmnožinu z n-prvkové množiny $(k, n \in \mathbb{N}_0)$.[23]

Kombinační čísla mají tyto vlastnosti:

$$\binom{n}{0} = \binom{n}{n} = \binom{0}{0} = 1, \quad \binom{n}{1} = n, \quad \binom{n}{n-k} = \binom{n}{k}, \quad \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}, \tag{5.2}$$

kde $k, n \in \mathbb{N}_0$ a $n \geq k$. [23]

Tyto uvedené vlastnosti kombinačních čísel je možno ilustrovat na tzv. **Pascalově trojúhelníku**:

$$\begin{array}{lccccccccccccc}
0 & & & & & & & \binom{0}{0} & & & & & & \\
1 & & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & \\
2 & & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
3 & & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & \\
4 & & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & \\
5 & & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & \\
\ldots & & & & & & & ... & & & & & & \\
\text{n} & \binom{n}{0} & & \binom{n}{1} & & & & ... & & & & \binom{n}{n-1} & & \binom{n}{n}
\end{array}$$

Po vyčíslení kombinačních čísel obdržíme:

$$\begin{array}{lccccccccccccc}
0 & & & & & & & 1 & & & & & & \\
1 & & & & & & 1 & & 1 & & & & & \\
2 & & & & & 1 & & 2 & & 1 & & & & \\
3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
4 & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
5 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
\ldots & & & & & & & ... & & & & & & \\
\text{n} & 1 & & n & & & & ... & & & & n & & 1
\end{array}$$

Nyní si již můžeme zavést **binomickou větu**:

Věta 5.2.1 ([23], s. 68)**.** Pro všechna reálná čísla a, b a každé přirozené číslo n platí

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k, \tag{5.3}$$

kde kombinační číslo

$$\binom{n}{k}$$

se nazývá **binomický koeficient**.

Binomickou větu si dokážeme dvěma způsoby – nejprve provedeme kombinatorický důkaz a poté důkaz pomocí matematické indukce.

Kombinatorický důkaz binomické věty

Budeme vycházet z rovnosti

$$(a+b)^n = (a+b)(a+b)...(a+b),$$

kde výraz $(a+b)$ násobíme mezi sebou n-krát, kde $a, b \in \mathbb{R}$ a $n \in \mathbb{N}_0$..

Myšlenku důkazu si nejdříve ukážeme pro $n = 3$:

- Máme-li vybrat a z každého činitele $(a+b)$ ze součinu $(a+b)^3$, získáme a^3. Výběr lze uskutečnit právě jedním způsobem.

- Vybereme-li b z libovolného jednoho činitele $(a+b)$, získáme b, a a ze zbývajících dvou činitelů, získáme a^2. Počet různých způsobů, jak vybrat jedno b z 3 činitelů je $\binom{3}{1} = 3$, zatímco a^2 můžeme vybrat pouze jedním způsobem. Proto výraz obsahující b má tvar $3a^2b$.

- Vybereme-li b z jakýchkoli dvou činitelů $(a+b)$, získáme b^2, a a ze zbývajícího činitele, získáme a. Počet různých způsobů, jak vybrat b^2 z tří činitelů je $\binom{3}{2} = 3$, zatímco a můžeme vybrat pouze jedním způsobem. Proto výraz obsahující b^2 má tvar $3ab^2$.

- Vybereme-li b ze všech tří činitelů $(a+b)$, obdržíme b^3. Výběr lze uskutečnit právě jedním způsobem.

Tedy platí

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 = \sum_{k=0}^{3} \binom{3}{k} a^{3-k} b^k.$$

Nyní myšlenku zobecníme pro n:

- Vybereme-li a z každého činitele $(a+b)$, obdržíme a^n, což můžeme provést právě jedním způsobem.

- Vybereme-li b z libovolného jednoho činitele $(a+b)$ a a ze zbývajících $(n-1)$ činitelů $(a+b)$. Počet různých způsobů, jak vybrat jedno b z n činitelů je $\binom{n}{1}$, zatímco a lze vybrat právě jedním způsobem. Proto výraz obsahující b má tvar $\binom{n}{1}a^{n-1}b$.

- Vybereme-li b z jakýchkoli dvou činitelů $(a+b)$ a a ze zbývajících $(n-2)$ činitelů. Počet různých způsobů, jak vybrat b^2 z n činitelů je $\binom{n}{2}$, zatímco a lze vybrat právě jedním způsobem. Proto výraz obsahující b^2 má tvar $\binom{n}{2}a^{n-2}b^2$.

  ...

- Vybereme-li b z libovolných i činitelů $(a+b)$, kde $i \in \mathbb{N}$, a a ze zbývajících $(n-i)$ činitelů. Počet různých způsobů, jak vybrat b^i z n činitelů je $\binom{n}{i}$, zatímco a můžeme vybrat právě jedním způsobem. Proto výraz obsahující b^i má tvar $\binom{n}{i}a^{n-i}b^i$.

  ...

- Vybereme-li b z každého činitele $(a+b)$ a tak obdržíme b^n, což lze provést právě jedním způsobem.

Tedy platí

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + ... + \binom{n}{i}a^{n-i}b^i + ... + \binom{n}{n-1}ab^{n-1} + b^n$$
$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k. \quad ([24], \text{s. } 14)$$

Důkaz binomické věty matematickou indukcí

1. Nejprve ověříme platnost pro $n = 0$:

   $$(a+b)^0 = 1,\ \sum_{k=0}^{0}\binom{0}{k}a^{0-k}b^k = \binom{0}{0}a^0b^0 = 1 \ \Rightarrow\ (a+b)^0 = \sum_{k=0}^{0}\binom{0}{k}a^{0-k}b^k.$$

   Pro $n = 0$ tedy rovnost platí.

2. Pro indukční krok budeme předpokládat, že věta platí pro exponent i. Dokážeme, že platí také pro $n = i + 1$:

   $$(a+b)^{i+1} = (a+b)(a+b)^i = (a+b)\sum_{k=0}^{i}\binom{i}{k}a^{i-k}b^k$$

   Chceme dokázat rovnost

   $$(a+b)\sum_{k=0}^{i}\binom{i}{k}a^{i-k}b^k = \sum_{k=0}^{i+1}\binom{i+1}{k}a^{i+1-k}b^k,$$

proto budeme levou stranu rovnosti upravovat tak, abychom získali tvar na pravé straně. Využijeme přitom platnost vztahů (5.2).

$$
\begin{aligned}
&(a+b)\sum_{k=0}^{i}\binom{i}{k}a^{i-k}b^k = \binom{i}{0}a^{i+1} + \binom{i}{0}a^ib + \binom{i}{1}a^ib + \binom{i}{1}a^{i-1}b^2 + \binom{i}{2}a^{i-1}b^2 + \\
&+ \ldots + \binom{i}{i-1}ab^i + \binom{i}{i}ab^i + \binom{i}{i}b^{i+1} = \\
&= \binom{i}{0}a^{i+1} + \binom{i+1}{1}a^ib + \binom{i+1}{2}a^{i-1}b^2 + \ldots + \binom{i+1}{i}ab^i + \binom{i}{i}b^{i+1} = \\
&= \binom{i+1}{0}a^{i+1} + \binom{i+1}{1}a^ib + \binom{i+1}{2}a^{i-1}b^2 + \ldots + \binom{i+1}{i}ab^i + \binom{i+1}{i+1}b^{i+1} = \\
&= \sum_{k=0}^{i+1}\binom{i+1}{k}a^{i+1-k}b^k,
\end{aligned}
$$

což je tvar, který jsme chtěli získat. Tím jsme dokázali platnost binomické věty pro všechna $n \in \mathbb{N}_0$.

Definice 5.2.2 ([23], s. 116)**. [Binomické rozdělení.]** Mějme n nezávislých pokusů, z nichž každý skončí buď zdarem s pravděpodobností p, nebo nezdarem s pravděpodobností q, kde $q = 1 - p$. Potom pravděpodobnost jevu A_k, že právě k pokusů bude zdařilých, je

$$P(A_k) = \binom{n}{k}p^kq^{n-k}, \ k = 0, 1, 2, ..., n. \tag{5.4}$$

Je zřejmé, že výrazy $\binom{n}{k}p^kq^{n-k}$ jsou součástí rozvoje $(p+q)^n$ binomické věty. Jevy „právě k pokusů bude zdařilých“ se pro $k = 0, 1, ..., n$ navzájem vylučují (neboť jsou nezávislé) a jeden z nich vždy nastává. Jejich sjednocení je proto jistý jev a tedy

$$\binom{n}{0}q^n + \binom{n}{1}pq^{n-1} + \ldots + \binom{n}{n}p^n = 1,$$

což také vyplývá z binomické věty, neboť $p+q = p+(1-p) = 1$, a proto také platí $(p+q)^n = 1$. Pravděpodobnosti $P(A_k)$ ze vzorce (5.4) tvoří **Binomické rozdělení**, neboli Bernoulliovo schéma.

Ukažme si, jak se vzorec (5.4) pro pravděpodobnosti tvořící binomické rozdělení odvozuje.

Odvození vzorce pro binomické rozdělení

Náhodný pokus má dva možné výsledky (zdar, nezdar) s pravděpodobnostmi p (zdar), q (nezdar). Provedeme jej $n-$krát po sobě tak, že jednotlivé pokusy jsou navzájem **nezávislé**.

Určíme pravděpodobnost možností:

1. Všechny dílčí pokusy skončí zdarem.
2. Nejdříve k pokusů skončí zdarem a zbývající pokusy skončí nezdarem.
3. Z n provedených pokusů skončí libovolných k zdarem.

Jedná se o nezávislé pokusy, proto budeme při sjednocování pravděpodobností dílčí jevů jejich pravděpodobnosti násobit.

1. Pro výsledek všechny pokusy skončí zdarem $Z \cdot Z \cdot Z \cdot ... \cdot Z$, budeme násobit pravděpodobnosti n nezávislých pokusů se shodnou pravděpodobností p:
$$P_1 = p \cdot p \cdot p \cdot ... \cdot p \cdot p = p^n.$$
2. Pro výsledek nejdříve k pokusů skončí zdarem a zbývající pokusy nezdarem $Z \cdot Z \cdot ... \cdot Z \cdot N \cdot N \cdot ... \cdot N$, násobíme $k-$krát pravděpodobnost zdaru p a $(n-k)-$krát pravděpodobnost q nezdaru:
$$P_2 = p \cdot p \cdot ... \cdot p \cdot q \cdot q \cdot ... \cdot q = p^k q^{n-k}.$$
3. Pro výsledek z n provedených pokusů skončí libovolných k zdarem, nám stačí libovolný výsledek, který vznikne přerovnáním výsledku z bodu 2., tj. všechny dílčí pokusy mají pravděpodobnost $p^k q^{n-k}$. Zbývá zjistit, kolik takových pravděpodobností je:

   - Vytváříme neuspořádané n-tice z k a $(n-k)$ prvků, dostáváme
   $$\frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}$$
   možností.
   - Nebo se lze stejného výsledku dobrat myšlenkou, že vybíráme z n umístění k takových, kdy bude výsledkem zdar: obdržíme
   $$\binom{n}{k}$$
   možností.

   Ve výsledku tedy $\binom{n}{k}$-krát násobíme pravděpodobnost $p^k q^{n-k}$, tj.
   $$P_3 = \binom{n}{k} p^k q^{n-k},$$
   kde $k = 0, 1, ..., n.$

Odvodili jsme vzorec pro pravděpodobnosti P tvořící binomické rozdělení (5.4).

5.3 Pravděpodobnost narozeného potomstva určité skladby pohlaví v jedné rodině

Nyní si objasníme, jak lze snadno pomocí binomické věty nebo binomického rozdělení určit pravděpodobnost narození potomstva určité skladby pohlaví narozených dětí v jedné rodině. Princip výpočtu si ukážeme rovnou na řešeném příkladu.

Příklad 5.3.1. Jaká je pravděpodobnost, že se v jedné rodině narodí v libovolném pořadí dva chlapci a dvě dívky (další děti už rodiče mít nebudou)?

Řešení:

Pravděpodobnost narození dívky je $P(d) = \frac{1}{2}$ a pravděpodobnost narození chlapce je také $P(ch) = \frac{1}{2}$. Vypočteme příklad více různými způsoby:

1. **Logicky (kombinatoricky) –** vypíšeme si všechny možnosti složení pohlaví 4 dětí pro 2 chlapce a 2 dívky a jejich pravděpodobnosti, které nakonec sečteme:

$$\begin{aligned} P(ch, ch, d, d) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16} \\ P(ch, d, d, ch) &= \frac{1}{16} \\ P(ch, d, ch, d) &= \frac{1}{16} \\ P(d, ch, d, ch) &= \frac{1}{16} \\ P(d, d, ch, ch) &= \frac{1}{16} \\ P(d, d, ch, ch) &= \frac{1}{16} \end{aligned}$$

   Existuje tedy 6 možností (pořadí), jak se ze 4 dětí můžou narodit 2 chlapci a 2 dívky. Celková pravděpodobnost narození 2 chlapců a 2 dívek v libovolném pořadí je $P(2ch, 2d) = = \frac{6}{16} = \frac{3}{8}$.

   Tento způsob je ale zbytečně zdlouhavý.

2. **Pomocí binomické věty –** do vzorce pro binomickou větu (5.4) dosadíme za n počet dětí v rodině, a bude reprezentovat pravděpodobnost narození chlapce a b pravděpodobnost narození dívky. Binomickou větu pro příslušný řád rozvineme a vybereme ten člen rozvoje, jehož mocniny odpovídají danému složení potomstva:

$$(a + b)^4 = a^4 + 4a^3b + \mathbf{6a^2b^2} + 4ab^3 + b^4$$

Součet $(a+b)$ znamená „narodí se chlapec s pravděpodobností a nebo se narodí dívka s pravděpodobností b". Rozvoj binomické věty pro $n=4$ nám tedy reprezentuje všechny možné pravděpodobnosti různých skladeb pohlaví u 4 dětí. Pravděpodobnost narození 2 dívek a 2 chlapců reprezentuje prostřední člen $6a^2b^2$. Tedy platí $P(2ch, 2d) = 6a^2b^2 =$ $= 6 \cdot (\frac{1}{2})^2 \cdot (\frac{1}{2})^2 = \frac{6}{16} = \frac{3}{8}$.

I tento způsob může být zdlouhavý, zvláště pokud je potomků v rodině více. Poté je vhodné využít část Pascalova trojúhelníku

$n=0$					1				
$n=1$				1		1			
$n=2$			1		2		1		
$n=3$		1		3		3		1	
$n=4$	1		4		**6**		4		1

díky kterému nemusíme rozvíjet binomickou větu a rychle zjistíme příslušný koeficient, ze kterého odvodíme celý člen **6**a^2b^2. Existuje však ještě lepší postup, který se používá v genetice nejčastěji.

3. **Pomocí binomického rozdělení** – do vzorce (5.4) dosadíme za n opět celkový počet dětí v rodině, p bude reprezentovat pravděpodobnost narození chlapce a q pravděpodobnost narození dívky, k je počet chlapců a $(n-k)$ je počet dívek. Proč tomu tak je? Jedná se totiž o n nezávislých pokusů, kdy se buď s pravděpodobností $\frac{1}{2}$ narodí chlapec, nebo s pravděpodobností $(1-\frac{1}{2})$ dívka.[1] Pravděpodobnost narození dvou chlapců a dvou dívek v této rodině je tedy:

$$P(2ch, 2d) = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{4!}{2!2!}\cdot\frac{1}{4}\cdot\frac{1}{4} = \frac{6}{16} = \frac{3}{8}.$$

Příklad 5.3.2. Jaká je pravděpodobnost, že se v jedné rodině narodí v libovolném pořadí jeden chlapec a tři dívky (další děti už rodiče mít nebudou)? K výpočtu použijte binomickou větu i binomické rozdělení.

[1]Tyto pravděpodobnosti jsou ovšem zaokrouhlené. Z dlouhodobých statistik vyplývá, že se s o něco větší pravděpodobností (přibližně o 15 setin) rodí chlapci. Viz např.: webové stránky Českého statistického úřadu https://www.czso.cz/csu/czso/4-obyvatelstvo-, tabulka 4-10. Bylo tomu tak vždy a to z důvodu toho, že jsou spermie nesoucí chromozom Y o něco málo pohyblivější.

Řešení:

- $n = 4$, $a = b = \frac{1}{2}$:

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + \mathbf{4ab^3} + b^4$$

$$P(1ch, 3d) = 4ab^3 = 4 \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

- $n = 4$, $p = q = \frac{1}{2}$, $k = 1$:

$$P(1ch, 3d) = \binom{4}{1}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 = 4 \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{4}$$

Pravděpodobnost narození jednoho chlapce a tří dívek v této rodině je $\frac{1}{4}$.

Příklad 5.3.3. Jaká je pravděpodobnost, že ze šesti dětí narozených v jedné rodině budou čtyři chlapci a dvě dívky? Určete v procentech.

Řešení:

$$P = \binom{6}{2}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^2 = \frac{6!}{4! \cdot 2!} \cdot \left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^2 = \frac{15}{64} \approx 23,4\ \%.$$

Pravděpodobnost narození šesti dětí ve složení čtyři chlapci a dvě dívky je přibližně $23,4\ \%$.

Příklad 5.3.4. Jaká je pravděpodobnost, že se v jedné rodině za sebou narodí čtyři chlapci? Určete v procentech.

Řešení:

$$P = \binom{4}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 = \frac{4!}{4! \cdot 0!} \cdot \left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 = \frac{1}{16} \approx 6,25\ \%.$$

Pravděpodobnost narození čtyř dětí ve složení samí chlapci je přibližně $6,25\ \%$.

U tohoto typu aplikace se setkáváme nejčastěji s úlohami na počty narozených dívek a chlapců, kde je pravděpodobnost početí každého pohlaví $\frac{1}{2}$. Stejný princip výpočtu lze však uplatnit i na podobných genetických příkladech, kde je pravděpodobnost jiná. Ukážeme si na vzorovém příkladu:

Příklad 5.3.5. Jaká je pravděpodobnost, že po zkřížení dvou heterozygotů (tj. jedinců s odlišnými alelami téhož genu) $Aa \times Aa$, kde A je dominantní alela pro červenou barvu květu a a recesivní alela pro bílou barvu květu, získáme osm červených a tři bílé květy?

Řešení: Uplatněním 2. Mendelova zákona o křížení heterozygotů (viz [27], s. 43 – 44) získáme fenotypový štěpný poměr potomků $3 : 1$, tedy s pravděpodobností $75\ \%$ vznikne rostlina s červeným květem (AA nebo Aa) a s pravděpodobností $25\ \%$ s bílým květem (aa). Proto má vzorec pro hledanou pravděpodobnost tvar

$$P = \frac{11!}{8! \cdot 3!} \cdot \left(\frac{3}{4}\right)^8 \left(\frac{1}{4}\right)^3 \approx 0,258 = 25,8\ \%.$$

Daná skladba potomstva vznikne s přibližnou pravděpodobností $25,8\ \%$.

5.4 Hardy-Weinbergova rovnováha ve velké panmiktické populaci

Populační genetika je odvětví genetiky zkoumající dědičnost a proměnlivost **populací**, tj. skupin jedinců téhož druhu, kteří jsou navzájem propojeni příbuzenskými vztahy vzniklými mj. vzájemným křížením.[25]

Panmiktická populace je taková populace, uvnitř níž se může jakákoliv samičí gameta (vajíčko) spojit se stejnou pravděpodobností s každou samčí gametou (spermií), tedy dochází k **náhodnému oplození** (tzv. panmixii).[25]

Model je záměrné zjednodušení složité reality, kdy se zanedbají nepodstatné faktory a zachovají naopak ty podstatné, aby bylo možné model zkoumat. Faktory populační genetiky jsou např. velikost populace, způsob oplození, mutace, migrace a přírodní výběr. Kvalita modelu závisí samozřejmě na tom, jak se moc blíží ideálu. V populační genetice se často používá **matematický model**, což je abstraktní model, který pomocí matematických vztahů[2] popisuje chování zkoumané soustavy nebo procesu.[26]

Hardy-Weinbergův (HW) model genotypových četností při náhodném oplození umožňuje v populaci určit genotypové četnosti pro gen se dvěma alelami. Předpoklady pro vytvoření tohoto matematického modelu jsou následující: organismy jsou diploidní[3] a rozmnožují se pohlavně, jejich oplození je náhodné, generace se nepřekrývají, jedná se o velkou populaci,

[2] Takovými vztahy jsou nejčastěji různé typy rovnic, jejich soustav nebo funkce.

[3] Tj. všechny jejich somatické buňky mají dvě sady autozomálních chromozomů.

migraci a mutaci lze zanedbat, na alely nepůsobí přírodní výběr.[26]

V populaci splňující tyto uvedené předpoklady platí tzv. Hardy-Weinbergův zákon:

Hardy-Weinbergův zákon o genetické rovnováze pro velkou panmiktickou populaci říká, že frekvence jednotlivých alel v panmiktické populaci zůstává konstantní, tj.

$$p + q = 1, \tag{5.5}$$

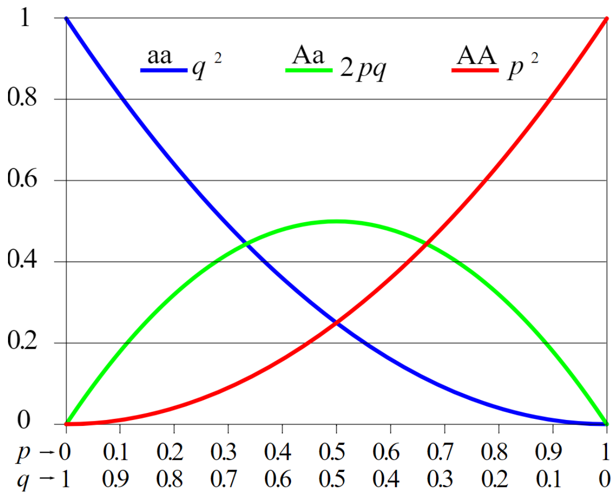
kde p je frekvence dominantní alely ozn. A, q frekvence recesivní alely ozn. a. Po jedné generaci náhodného oplození získáváme v populaci

$$(p + q)^2 = p^2 + 2pq + q^2 = 1\ (= 100\ \%), \tag{5.6}$$

kde p^2 je frekvence dominantního homozygota AA, $2pq$ je frekvence heterozygota Aa, q^2 je frekvence recesivního homozygota aa.[27]

Odvození H-W zákona lze najít např. v [25] s. 351 – 356.

Pokud by došlo k vychýlení frekvence genotypů ze své stability, ustálí se nová rovnováha v následující generaci. Tento jednoduchý matematický model se dobře používá při výzkumu genofondů různých populací, lidskou populaci nevyjímaje.[27]



Obrázek 5.1: Hardy-Weinbergův princip pro dvě alely. Na ose x je frekvence alel, na ose y frekvence genotypů. (Převzato z: [28])

Velikou výhodou u populace v genetické rovnováze je, že můžeme frekvence alel a genotypů určit pouze na základě známé frekvence jednoho genotypu (nejčastěji recesivního homozygota).

Na následujících příkladech si objasníme H-W princip. Nejprve si ukážeme nejjednodušší příklad, kdy se z frekvence recesivních homozygotů pomocí vzorce (5.6) snadno určí frekvence obou alel a zbývajících genotypů.

Příklad 5.4.1. V rozsáhlé panmiktické populaci bylo zjištěno $9~\%$ recesivních homozygotů aa. Zjistěte v této populaci:

a) frekvenci obou alel příslušného genu,

b) frekvenci dominantních homozygotů,

c) frekvenci heterozygotů.

Řešení:

a) frekvence recesivního homozygota aa ... $q^2 = 0,09$

   frekvence recesivní alely a ... $q = \sqrt{q^2} = \sqrt{0,09} = 0,3 = 30~\%$

   frekvence dominantní alely A ... $p = 1 - q = 1 - 0,3 = 0,7 = 70~\%$

b) frekvence dominantních homozygotů AA ... $p^2 = 0,7^2 = 0,49 = 49~\%$

c) frekvence heterozygotů Aa ... $2pq = 2 \cdot 0,7 \cdot 0,3 = 0,42 = 42~\%$

Zkouška: $p^2 + 2pq + q^2 = 0,49 + 0,42 + 0,09 = 1$.

Frekvence dominantní alely je $70~\%$, frekvence recesivní alely je $30~\%$. Frekvence dominantních homozygotů je $49~\%$. Frekvence heterozygotů je $42~\%$.

Příklad 5.4.2. Cystická fibrosa je recesivní choroba vázaná na autozomy (nepohlavní chromozomy)[4], která se projevuje chronickým onemocněním dýchacích plic. Cystická fibróza v naší populaci postihuje jednoho z 2 500 jedinců (1 : 2 500). Kolik je mezi námi přenašečů? [5]

Řešení:

Recesivní choroba vázaná na autozomy je taková choroba, kde gen pro danou nemoc leží na autozomu a zároveň je podmíněna recesivní alelou a, která zapříčiňuje onemocnění, vliv dominantní alely A však zcela potlačí její účinek. Proto daný jedinec s jednou recesivní alelou

[4] V jádře každé naší tělní buňky je 23 párů chromozomů – z toho 22 párů jsou autozomy a jeden pár jsou pohlavní chromozomy, tzv. gonozomy, u muže XY, u ženy XX.

[5] informace převzaty z: https://cs.wikipedia.org/wiki/Cystická_fibróza

a jednou dominantní (heterozygot) je zdravý, recesivní alelu ovšem přenáší do dalších generací, proto jej označujeme jako „přenašeč“. U člověka se projeví nemoc pouze v případě, že jsou obě jeho alely v genu pro cystickou fibrózu recesivní. Existují tedy možnosti:

genotyp	**fenotyp**
AA	zdravý jedinec
Aa	zdravý jedinec, ale přenašeč
aa	nemocný jedinec

- frekvence aa ... $1 : 2500$ ... $q^2 = \dfrac{1}{2500}$
- frekvence a ... $q = \sqrt{\dfrac{1}{2500}} = \dfrac{1}{50}$
- frekvence A ... $p = 1 - q = \dfrac{49}{50}$
- frekvence Aa ... $2pq = 2 \cdot \dfrac{1}{50} \cdot \dfrac{49}{50} \approx \dfrac{1}{25}$ ... $1 : 25$

Zkouška: $p^2 + 2pq + q^2 = \left(\frac{49}{50}\right)^2 + \frac{1}{25} + \frac{1}{2500} = 1$
Z 25 lidí je jeden přenašeč cystické fibrózy (1 : 25).

Příklad 5.4.3. V populaci České republiky je 84 % obyvatel Rh pozitivních (Rh$^+$) a 16 % obyvatel je Rh negativních (Rh$^-$). Rh-faktor je autozomálně dominantní. Jedná se o krevně skupinový systém, podobně jako krevní skupiny A, B, AB a 0. Určete četnost alely (A) pro Rh faktor a četnosti všech genotypů.[6]

Řešení:
Rh-faktor je autozomálně dominantní, což znamená, že sledovaný gen pro Rh-faktor leží na autozomech a zároveň že jedinec s alespoň jednou dominantní alelou A bude Rh pozitivní (Rh$^+$):

genotyp	**fenotyp**
AA	Rh$^+$
Aa	Rh$^+$
aa	Rh$^-$

[6]informace převzaty z: https://cs.wikipedia.org/wiki/Krevní_skupina♯Rhesus_(Rh)_faktor

- frekvence aa ... $q^2 = 16\ \% = 0,16$
- frekvence a ... $q = \sqrt{0,16} = 0,4 = 40\ \%$
- frekvence A ... $p = 1 - q = 0,6 = 60\ \%$
- frekvence AA ... $p^2 = 0,6^2 = 0,36 = 36\ \%$
- frekvence Aa ... $2pq = 2 \cdot 0,6 \cdot 0,4 = 48\ \%$

Zkouška: $p^2 + 2pq + q^2 = 0,36 + 0,48 + 0,16 = 1$
Četnost alely pro Rh-faktor je 60 %. Četnost jednotlivých genotypů je: dominantní homozygot 36 %, heterozygot 48 %, recesivní homozygot 16 %.

Příklad 5.4.4. Rostlinu štírovník růžkatý (*Lotus corniculatus*) chrání kyanogenní glykosid před hmyzími škůdci a dokonce i před pasoucím se dobytkem. Tento glykosid je přítomen, pokud se v genotypu rostliny vyskytuje alespoň jedna dominantní alela. Zkoumaná populace štírovníku se skládá z 77 rostlin, které mají kyanogenní glykosid a 56 rostlin, které ho nemají. Jaká je frekvence dominantní alely, která způsobuje přítomnost kyanogenního glykosidu, v této populaci?[7]

Řešení:
Danou dominantní alelu označme opět A. Víme, že je celkem v populaci 133 rostlin, z toho 77 má glykosid (genotyp AA nebo Aa) a 56 nemá glykosid (genotyp aa).

- frekvence aa ... $q^2 = \dfrac{56}{133} \approx 0,42$
- frekvence a ... $q = \sqrt{0,42} \approx 0,65$
- frekvence A ... $p = 1 - q = 0,35 = 35\ \%$

Zkouška: $p^2 + 2pq + q^2 = (0,35)^2 + 2 \cdot 0,65 \cdot 0,35 + 0,42 = 1$
Frekvence dané dominantní alely v populaci štírovníku růžkatého je přibližně $35\ \%$.

[7] informace převzaty z: https://cs.wikipedia.org/wiki/Štírovník_růžkatý

5.5 Pracovní list

Téma: Aplikace binomické věty a binomického rozdělení v genetice

Jméno a třída: Datum:

Příklad 1: Jaká je pravděpodobnost, že ze tří dětí narozených v jedné rodině budou dva chlapci a jedna dívka? Vypočítej pomocí binomické věty i binomického rozdělení.

Příklad 2: Jaká je pravděpodobnost, že ze sedmi dětí narozených v jedné rodině budou tři chlapci a čtyři dívky? Urči v procentech.

Příklad 3: V nejznámějším románu anglické spisovatelky Jane Austenové s názvem Pýcha a předsudek měla paní Bennetová pět dcer, které se snažila za každou cenu a co nejvýhodněji provdat. Další děti neměla. Jaká je pravděpodobnost, že se v jedné rodině narodí po sobě pět dcer?

Příklad č. 4 V rozsáhlé panmiktické populaci bylo zjištěno $16\ \%$ recesivních homozygotů. Zjisti v této populaci

a) frekvenci obou alel příslušného genu,

b) frekvenci dominantních homozygotů,

c) frekvenci heterozygotů.

Příklad č. 5 Fenylketonurie, porucha metabolismu aminokyseliny fenylalaninu, je autozomálně recesivní choroba, která se v naší populaci vyskytuje přibližně u jednoho z 10 000 jedinců . Kolik je mezi námi přenašečů?[8]

Příklad č. 6 Tay-Sachsova choroba je autozomálně recesivní porucha. Jedná se o velmi vzácnou metabolickou chorobu. Mezi židy Aškenazi (středoevropská židovská větev) je nemoc přítomna u jednoho člověka z 3600. Pokud se populace Aškenazi páruje náhodně pro Tay-Sachs gen, jakou část populace tvoří přenašeči a jakou část zdraví jedinci nepřenašeči?[9]

[8]informace převzaty z: https://cs.wikipedia.org/wiki/Fenylketonurie

[9]informace převzaty z: https://www.wikiskripta.eu/w/Tay-Sachsova_choroba

Řešení pracovního listu:

Příklad 1: Výpočet podle

1. binomické věty $(a+b)^n$, kde $n = 3$ a $a = b = \frac{1}{2}$: pro nás vyhovující člen rozvoje binomické věty, do kterého dosadíme pravděpodobnost narození daných pohlaví, je $3a^2b$, tedy platí $P = 3\left(\frac{1}{2}\right)^2 \frac{1}{2} = \frac{3}{8}$,
2. binomického rozdělení $\binom{n}{k}p^k q^{n-k}$, kde $n = 3, p = q = \frac{1}{2}, k = 2$: $P = \binom{3}{2}\left(\frac{1}{2}\right)^2 \frac{1}{2} =$
$= \frac{3}{8}$.

Pravděpodobnost narození dvou chlapců a jedné dívky v dané rodině je $\frac{3}{8}$.

Příklad 2: Použijeme vzorec pro binomické rozdělení $\binom{n}{k}p^k q^{n-k}$, kde $n = 7$, $k = 3$ a $p = q = \frac{1}{2}$. Tedy $\binom{7}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^4 = \frac{35}{128} \approx 27,3\ \%$. Pravděpodobnost narození tří chlapců a čtyř dívek v dané rodině je přibližně $27,3\ \%$.

Příklad 3: Opět použijeme vzorec pro binomické rozdělení $\binom{n}{k}p^k q^{n-k}$, kde $n = 5$, $k = 0$ a $p = q = \frac{1}{2}$. Potom $\binom{5}{0}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 = \frac{1}{32} \approx 3,13\ \%$. Pravděpodobnost narození pěti dívek v rodině Bennetových je přibližně $3,13\ \%$.

Příklad 4: Využijeme Hardy-Weinbergova zákona o konstantní frekvenci alel v panmiktické populaci: $p+q = 1$ a $(p+q)^2 = 1$. Tedy platí $q = \sqrt{0,16} = 0,4 = 40\ \%$, potom $p = 1-q =$
$= 1 - 0,4 = 0,6 = 60\ \%$, $p^2 = 0,36 = 36\ \%$ a $2pq = 0,48 = 48\ \%$. a) Frekvence alel je $40\ \%$ a $60\ \%$. b) Frekvence dominantních homozygotů je $36\ \%$. c) Frekvence heterozygotů je $48\ \%$.

Příklad 5: Využijeme opět Hardy-Weinbergova zákona. Frekvence 1 z 10 000 znamená, že frekvence $q^2 = \frac{1}{10000} \Rightarrow q = \sqrt{\frac{1}{10000}} = \frac{1}{100}$, $p = 1 - q = \frac{99}{100}$ a tedy $2pq = \frac{99}{5000} \approx \frac{1}{50}$. Přibližně každý 50. člověk v naší populaci je přenašečem fenylketonurie.

Příklad 6: Využijeme opět Hardy-Weinbergova zákona. Frekvence 1 z 3600 znamená, že frekvence $q^2 = \frac{1}{3600} \Rightarrow q = \sqrt{\frac{1}{3600}} = \frac{1}{60}$, $p = 1 - q = \frac{59}{60} \Rightarrow p^2 = \frac{3481}{3600} \approx 96,7\ \%$ a tedy $2pq = \frac{59}{1800} \approx 3,3\ \%$. Frekvence přenašeče Tay-Sachsovy choroba je v dané populaci přibližně $3,3\ \%$. Frekvence zdravého člověka nepřenašeče je přibližně $96,7\ \%$.

6. Závěr

Cílem mojí diplomové práce bylo propojit vybrané části dvou středoškolských předmětů – matematiky a biologie, a vytvořit tak inspirativní popř. studijní materiál pro učitele biologie, kteří jako druhou aprobaci nestudovali matematiku, učitelům matematiky, kteří nestudovali biologii, a studentům středních škol. Zároveň jsem si kladla za cíl poukázat na široké využití matematiky v biologii a lékařství.

Vybrala jsem tyto čtyři části:

1. podmíněnou pravděpodobnost a její využití při ověřování kvality lékařských diagnostických testů,

2. aplikaci exponenciální funkce ve vybraných populačních modelech,

3. uplatnění logických funkcí při matematickém modelování neuronu,

4. využití binomické věty a binomického rozdělení v genetice.

Pro každou z těchto aplikací jsem zpracovala matematický a biologický teoretický základ, připravila komentované příklady a vytvořila pracovní listy pro studenty včetně autorských řešení. První dvě aplikace jsem otestovala ve výuce středoškolské matematiky. Většinu uvedených příkladů jsem sama vymyslela.

První aplikaci z diplomové práce jsem také představila na matematické konferenci Užití počítačů ve výuce matematiky 2017 a zpracovala ve článku s názvem Aplikace podmíněné pravděpodobnosti v lékařské diagnostice. Článek vyšel ve sborníku této konference a je v celém rozsahu k nalezení v příloze diplomové práce nebo na adrese http://home.pf.jcu.cz/∼upvm/2017/wp-content/uploads/2018/02/Sbornik_ UPVM_2017.pdf.

První dvě kapitoly včetně pracovních listů jsem otestovala v praxi odučením v hodinách středoškolské matematiky. Studenty jsem také požádala o vyplnění krátkých dotazníků, které mi sloužily jako zpětná reflexe. Tyto dotazníky a jejich vyhodnocení spolu s popisem průběhu odučených hodin jsou k nalezení na konci daných kapitol.

Přínos diplomové práce shledávám především v možnosti ukázat studentům středních škol úzké propojení matematiky a biologie, a tudíž i význam matematiky na poli přírodních věd.

Dále také v podání inspirace středoškolským učitelům matematiky a biologie, jak by mohli své hodiny obohatit a zaujmout studenty pro dané téma.

K napsání diplomové práce byl použit sázecí systém LaTeX, pro vykreslení grafů pak programy GeoGebra a MS Excel.

Literatura

[1] STUDENÁ, Lucie. Aplikace podmíněné pravděpodobnosti v lékařské diagnostice. *In: Sborník příspěvků 8. konference Užití počítačů ve výuce matematiky, 9. – 11. listopadu 2017, České Budějovice* [online]. Společnost učitelů matematiky, JČMF, Katedra matematiky Pedagogické fakulty, Jihočeská univerzita v Č. Budějovicích, 2017, s. 132 - 146 [cit. 2018-03-09].

Dostupné z: http://home.pf.jcu.cz/~upvm/2017/wp-content/uploads/2018/02/Sbornik_UPVM_2017.pdf

[2] Dupač V., Hájek J. *Pravděpodobnost ve vědě a technice*. 1. vydání, Academia, 1962

[3] Hátle, J., Kahounová, J.: *Úvod do teorie pravděpodobnosti*. SNTL/ALFA, Praha, 1987.

[4] Podmíněná pravděpodobnost a Bayesův vzorec. *Matematická biologie: E-learningová učebnice* [online]. [cit. 2017-03-15].

Dostupné z: http://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat–biostatistika-pro-matematickou-biologii–vztah-pravdepodobnosti-statistiky-a-biostatistiky–podminena-pravdepodobnost-a-bayesuv-vzorec

[5] Senzitivita, specificita a prediktivní hodnoty. *Matematická biologie: E-learningová učebnice* [online]. [cit. 2017-03-15].

Dostupné z: http://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat–biostatistika-pro-matematickou-biologii–vztah-pravdepodobnosti-statistiky-a-biostatistiky–senzitivita-specificita-a-prediktivni-hodnoty

[6] ZVÁROVÁ J. *Základy statistiky pro biomedicínské obory.* Nakladatelství Karolinum, Praha, 2004.

[7] Kapitola 3 Diagnostika. *Medicína založená na důkazu* [online]. [cit. 2017-03-26].

Dostupné z: http://www.khshk.cz/e-learning/kurs1b/kapitola_3__diagnostika. html

[8] TOWNSEND, Colin R., Michael BEGON a John L. HARPER. *Základy ekologie.* V Olomouci: Univerzita Palackého, 2010. ISBN 978-80-244-2478-1.

[9] TKADLEC, Emil. *Populační ekologie: struktura, růst a dynamika populací.* 2. vyd. Olomouc: Univerzita Palackého v Olomouci, 2013. ISBN 978-80-244-3385-1.

[10] PENGRA, Bruce. *One Planet, How Many People?* A Review of Earth's Carrying Capacity. United Nations Environment Program, 2012 [online]. [cit. 2018-04-04].

Dostupné z: https://na.unep.net/geas/archive/pdfs/geas_jun_12_carrying_capacity.pdf

[11] ODVÁRKO, Oldřich. *Matematika pro gymnázia: funkce.* 3. upr. vyd. Praha: Prometheus, 1999. Učebnice pro střední školy. ISBN 80-7196-164-7.

[12] HŘEBÍČEK, Jiří, Zdeněk POSPÍŠIL a Jaroslav URBÁNEK. Úvod do matematického modelování s

[13] BODINE, Erin N., Suzanne LENHART a Luis J. GROSS. *Mathematics for the Life Sciences.* Princeton a Oxford: Princeton University Press, 2014. 640 stran. ISBN: 978-0-691-15072-7.

[14] MACHOVÁ, Jitka. *Biologie člověka pro učitele.* Druhé vydání. Praha: Univerzita Karlova v Praze, nakladatelství Karolinum, 2016. ISBN 978-80-246-3357-2.

[15] TROJAN, Stanislav a Michael SCHREIBER. *Knižní atlas biologie člověka: + 430 modelových otázek k přijímacím zkouškám na medicínu + 100 obrazových podkladů k opakování a procvičování.* 2., upr. vyd. Praha: Scientia, 2007. ISBN 978-80-86960-11-1.

[16] ROUGIER, Nicolas. Biological neuron schema. In: Wikimedia Commons [online]. Licence Creative Commons [cit. 2017-12-04]. Dostupné z: https://commons.wikimedia.org/wiki/File:Neuron-figure-notext.svg?uselang=cs

[17] VOLNÁ, Eva. *Neuronové sítě.* Ostrava: Ostravská univerzita v Ostravě, 2013. ISBN 978-80-7464-329-3.

[18] Šíma, J., Neruda, J. *Teoretické otázky neuronových sítí.* Praha: Matfyzpress, 1996.

[19] KŘIVÝ, Ivan. *Diskrétní matematika pro informatiky.* [Učební text.]. Ostrava: Ostravská univerzita v Ostravě, 2012.

[20] SOCHOR, Antonín. *Klasická matematická logika.* V Praze: Karolinum, 2001. ISBN 80-246-0218-0.

[21] KOPKA, Jan a obr. narýsoval Vlastimil MACHÁČEK. *Svazy a Booleovy algebry:* Celostát. vysokošk. učeb. pro stud. matematicko-fyz., přírodověd., ped. fak. stud. oboru učitelství všeobecně vzděl. předmětů aprobačního předmětu matem. Ústí nad Labem: Univ. J. E. Purkyně, 1991. ISBN 8070440252.

[22] Matematická biologie, E-learningová učebnice: Matematické modely v biologii: Matematické modely neuronu [online]. [cit. 2017-06-13].

Dostupné z: http://portal.matematickabiologie.cz/index.php?pg=analyza-a-modelovani-dynamickych-biologickych-dat–matematicke-modely-v-biologii–matematicke-modely-neuronu

[23] CALDA, Emil a Václav DUPAČ. *Matematika pro gymnázia: kombinatorika, pravděpodobnost, statistika.* 4. upr. vyd. Praha: Prometheus, 1999. Učebnice pro střední školy. ISBN 8071961477.

[24] DR EVANS, Michael AMSI. *The binomial theorem: A guide for teachers – Years 11 and 12* [online]. Australian Mathematical Sciences Institute, The University of Melbourne: Education Services Australia, 2013 [cit. 2017-07-24]. Dostupné z: https://amsi.org.au/ESA_Senior_Years/PDF/Thebinomialtheorem1c.pdf

[25] NEČÁSEK, Jan. *Obecná genetika: vysokošk. učebnice. 2. vyd.* Ilustroval Otakar PROCHÁZKA, ilustroval Miroslava JAKEŠOVÁ. Praha: SPN, 1984. Našim učitelům.

[26] RELICHOVÁ, Jiřina. *Genetika populací.* Brno: Masarykova univerzita, 1997. ISBN 80-210-1542-X.

[27] ŠMARDA, Jan. *Genetika pro gymnázia.* Praha: Fortuna, 2003. ISBN 80-7168-851-7.

[28] JOHNUNIQ. Hardy–Weinberg principle for two alleles. In: Wikimedia Commons [online]. Licence Creative Commons [cit. 2018-02-04]. Dostupné z: https://commons.wikimedia.org/wiki/File:Hardy-Weinberg.svg

A. Článek ze sborníku konference Užití počítačů ve výuce matematiky 2017

APLIKACE PODMÍNĚNÉ PRAVDĚPODOBNOSTI V LÉKAŘSKÉ DIAGNOSTICE

Bc. Lucie Studená

Ústav matematiky a biomatematiky, PřF, Jihočeská univerzita v Českých Budějovicích

Abstrakt: Článek se věnuje aplikaci podmíněné pravděpodobnosti a Bayesova vzorce v lékařské diagnostice. Na několika příkladech je ukázáno, jak se statisticky ověřuje správnost diagnostických testů a je objasněno, jaký význam má specificita a senzitivita lékařského testu. Toto praktické využití teorie pravděpodobnosti v lékařství může být uplatněno jak ve výuce matematiky, tak ve výuce biologie.

Klíčová slova: podmíněná pravděpodobnost, Bayesův vzorec, lékařská diagnostika, diagnostický test

Application of conditional probability in medical diagnostics

Abstract: The paper deals with the application of conditional probability in medical diagnostics. The statistical verification of diagnostic tests accuracy is illustrated with a few examples and it clarifies the meaning of the sensitivity and the specificity of medical tests. This practical application of probability theory can be effectively utilized in the teaching of mathematics and of biology.

Key words: conditional probability, Bayes′ formula, medical diagnostics, diagnostic test

A.1 Úvod

Podmíněná pravděpodobnost a Bayesův vzorec nachází své uplatnění v populačních etiologických studiích (studiích o příčinách a původech nemocí) a v některých matematických modelech diagnostického, terapeutického či prognostického lékařského rozhodování, např. tedy v případech, kdy je potřeba statisticky vyhodnotit kvalitu lékařského diagnostického testu. Statistickým vyhodnocováním lékařských diagnostických testů se zabývá lékařská diagnostika.

Tématem tohoto článku je právě aplikace podmíněné pravděpodobnosti v lékařské diagnostice. Jedná se v podstatě o jednu z kapitol mojí diplomové práce Aplikace matematických znalostí při výuce biologie, v níž propojuji oba obory, které studuji, ve snaze demonstrovat široké uplatnění a důležitost matematického aparátu na poli biologie a lékařství.

Cílem tohoto článku je především poskytnout učitelům středních škol návod jak studentům přiblížit teorii podmíněné pravděpodobnosti a Bayesův vzorec pomocí zajímavé aplikace; dále jak studentům ukázat důležitost teorie podmíněné pravděpodobnosti v lékařství; a v neposlední řadě také jak obohatit výuku a zaujmout studenty pomocí vysvětlení principu vyhodnocování diagnostických testů na zajímavých příkladech z běžného života.

Čtenář se dočte mnohé o diagnostických testech (jaké jsou druhy testů, kde se s nimi můžeme setkat, na jakých principech určité testy fungují apod.), o jejich ukazatelích správnosti, zejm. senzitivitě a specificitě, o principu statistického testování diagnostických testů, o vlivu prevalence na výsledek diagnostického testu aj.

Výhodou této aplikace je její využitelnost jak při výuce středoškolské matematiky, tak i při výuce středoškolské biologie. Učitelé matematiky mohou prostřednictvím aplikace studentům ukázat praktické využití poněkud neoblíbeného a často opomíjeného tématu podmíněné pravděpodobnosti. Učitelé biologie mají možnost studenty seznámit s diagnostickými testy, což jsou sice znalosti nad rámec běžné výuky, avšak mohou být pro studenty zajímavým zpestřením. Setkat se v běžném životě s různými diagnostickými testy, ať už s těmi zakoupitelnými v lékárně (těhotenský test, test na alkohol, test na jištění množství vitamínu B_{12} v těle apod.), nebo u lékaře (např. jaterní test na přítomnost fibrózy, zátěžové kardio testy apod.), totiž není nijak nevšední.

Je pouze na uvážení každého učitele, v jakém rozsahu a obtížnosti aplikaci ve svých hodinách využije a zda-li ji zařadí do běžné výuky, nebo ji představí jen vybrané skupině přírodovědně nadaných žáků, např. v rámci semináře. Učitelé se také mohou nechat inspirovat přiloženým návrhem pracovního listu se zajímavými příklady, díky kterým si studenti procvičí získané znalosti o diagnostických testech a naučí se také orientovat v jednoduchých tabulkách. K usnadnění výpočtů hodnot ze zadaných tabulek lze také využít MS Exel.

A.2 Podmíněná pravděpodobnost a Bayesův vzorec

Než si objasníme princip vyhodnocování kvality diagnostického testu, připomeneme si nezbytné znalosti z teorie podmíněné pravděpodobnosti.

„*Podmíněná pravděpodobnost značí, že k pravděpodobnosti nastání určitého jevu přidáme ještě nějakou podmínku.*“

Číslo $P(A)$ značí pravděpodobnost jevu A vzhledem k pevně danému souboru základních podmínek, po jejichž uskutečnění (provedení náhodného pokusu) nastane jev A s pravděpodobností $P(A)$. Pokud k těmto základním podmínkám připojíme ještě další podmínku, že např. nastal jev B ($P(B) > 0$), některé pravděpodobnosti se stanou jistými (např. jev B), některé nemožnými (např. každý jev neslučitelný s B) a ostatní jevy očekáváme se změněnými pravděpodobnostmi. Těmto pravděpodobnostem říkáme podmíněné pravděpodobnosti. [3],[4]

Definice A.2.1. ([4], s. 14)[**Podmíněná pravděpodobnost.**] Podmíněnou pravděpodobnost jevu A vzhledem k jevu B označíme $P(A|B)$ a definujeme ji jako:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \; P(B) > 0, \tag{A.1}$$

Uveďme si následující příklad. Mějme dva různé jevy A a B, jev A značí, že je určitý pacient určen jako pozitivní a jev B značí, že pacient skutečně trpí danou nemocí. $P(A)$ je pravděpodobnost, že je pacient určen jako pozitivní, $P(B)$ je pravděpodobnost, že pacient skutečně trpí nemocí. Podmíněná pravděpodobnost, že je *pacient určen jako pozitivní za předpokladu, že je skutečně nemocný* by se spočítala jako $P(A|B)$, kde $P(B) > 0$, dle vzorce (5.6) a pravděpodobnost, že je *pacient určen jako pozitivní za předpokladu, že nemocný není* by se dle stejného vzorce spočítala jako $P(A|B')$, kde $P(B') > 0$ a jev B' (pacient není nemocný) je jev opačný k jevu B, a tedy platí $P(B') = 1 - P(B)$.

S podmíněnou pravděpodobností souvisí pojem nezávislost jevů, tzn. že nastání jednoho jevu nemá vliv na nastání nebo nenastání druhého jevu.

Věta A.2.1. ([4], s. 15)**[Věta o násobení pravděpodobností.]** Nechť A, B jsou dva nezávislé jevy, pak pro pravděpodobnost jejich současného nastání platí

$$P(A \cap B) = P(A) \cdot P(B). \tag{A.2}$$

Jsou-li tedy jevy A, B nezávislé, pak dosazením rovnosti (A.2) do vzorce pro výpočet podmíněné pravděpodobnosti (A.1) získáme:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

Platí tedy, že podmíněná pravděpodobnost vzhledem k nezávislému jevu se rovná nepodmíněné pravděpodobnosti. [4]

Ze vzorce pro podmíněnou pravděpodobnost (A.1) si vyjádříme $P(A \cap B)$ následujícím způsobem:

$$P(A \cap B) = P(A|B)P(B). \tag{A.3}$$

Pokud budeme chtít vyjádřit pravděpodobnost jevu B za podmínky nastoupení jevu A, zaměníme ve vzorci pro podmíněnou pravděpodobnost (A.1) jevy A a B

$$P(B|A) = \frac{P(A \cap B)}{P(A)},\ P(A) > 0. \tag{A.4}$$

Rovnost (2.3) dosadíme do vzorce (A.4)

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)},\ P(B) > 0,$$ [5]

čímž jsme získali tzv. **Bayesův vzorec** pro dvojici jevů, který udává jak podmíněná pravděpodobnost nějakého jevu souvisí s opačnou podmíněnou pravděpodobností.

Věta A.2.2. ([3], s. 31)**[Bayesův vzorec.]** Mějme dva náhodné jevy A a B s pravděpodobnostmi $P(A)$ a $P(B)$, $P(B) > 0$, potom platí

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}, \tag{A.5}$$

kde $P(A|B)$ je podmíněná pravděpodobnost jevu A za předpokladu, že nastal jev B.

Věta A.2.3. ([3], s. 29)**[O úplné pravděpodobnosti.]** Nechť B, C jsou dva jevy, které se navzájem vylučují, přičemž jeden z nich nutně nastává. Potom pro libovolný jev A platí $A = A \cap B \cup A \cap C$, pak pravděpodobnost $P(A)$ lze vyjádřit

$$P(A) = P(B)P(A|B) + P(C)P(A|C). \tag{A.6}$$

Bayesův vzorec (A.5) lze dále upravit dosazením vzorce (A.6) do jmenovatele za $P(A)$:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(B)P(A|B) + P(C)P(A|C)}.\text{[5]} \tag{A.7}$$

A.3 Lékařská diagnostika

A.3.1 Princip statistického vyhodnocování diagnostických testů

V úvodu článku jsme se dozvěděli, že podmíněná pravděpodobnost se v lékařství uplatňuje při statistickém **vyhodnocování správnosti diagnostických testů**. Diagnostický test je vlastně určitá speciální metoda pomáhající s určitou pravděpodobností zjistit, zda pacient má, resp. nemá, zjišťovanou nemoc nebo určitou látku v těle.

Test v lékařské diagnostice může představovat nejen určitou laboratorní metodu, např. histologické vyšetření, ale i např. ultrazvukové vyšetření. Mezi diagnostické testy patří také testy volně zakoupitelné v lékárně jako např. těhotenský test nebo test potravinové intolerance. Na trhu existuje takovýchto různých testů nepřeberné množství
(viz např. www.prozdravi.cz/diagnosticke-testy).

Požadavky na diagnostické testy jsou, aby byly jednoduché, snadno a rychle proveditelné, finančně příliš nenákladné a také neškodné a bezbolestné. [2]

Nyní si ukážeme, jak se provádí statistické vyhodnocování správnosti diagnostického testu a co vše je k tomu potřeba znát.

Každý diagnostický test má dva možné výsledky, jedná se vlastně o dvojici opačných a tedy i navzájem neslučitelných jevů. Nazýváme je:

- **pozitivní výsledek**, ozn. A^+,
- **negativní výsledek**, ozn. A^-,

Dále je důležitá také skutečná přítomnost nemoci, což jsou opět dva opačné navzájem neslučitelné jevy nazývané:

- **nemoc je přítomna**, ozn. H^+,
- **nemoc není přítomna**, ozn. H^-. [2]

Diagnostické schopnosti testu jsou při vyhodnocování prověřovány proti skutečnému stavu testovaných osob, kde se proti sobě srovnávají právě pozitivní (resp. negativní) výsledky testu a prokazatelně zjištěná přítomnost (resp. nepřítomnost) nemoci nebo látky. Pro takovéto vyhodnocování je zavedeno několik ukazatelů správnosti (validity) představujících číselné ohodnocení testu ve vztahu k jeho chybovosti. Představíme si čtyři nejzákladnější z nich. [2]

První dva ukazatele správnosti diagnostických testů jsou tzv. *senzitivita testu* (schopnost testu rozpoznat nemocné osoby) a *specificita testu* (schopnost testu rozpoznat osoby bez nemoci). Tyto ukazatele definujeme pomocí podmíněné pravděpodobnosti:

- **senzitivita testu** je pravděpodobnost, že test bude pozitivní, když je osoba skutečně nemocná, tj. $P(A^+|H^+)$,
- **specificita testu** je pravděpodobnost, že test bude negativní, když je osoba zdravá, tj. $P(A^-|H^-)$. [2]

S pojmy senzitivita a specificita se můžete běžně setkat v příbalovém letáku u volně zakoupitelných testů, nebo v popisu testu v internetové lékárně
Druhé dva ukazatele správnosti diagnostických testů jsou tzv. *prediktivní hodnoty*, které stejně jako předchozí ukazatele definujeme pomocí podmíněné pravděpodobnosti:

- **prediktivní hodnota pozitivního testu** je pravděpodobnost, že osoba skutečně je nemocná, když test vyjde pozitivní, tj. $P(H^+|A^+)$,
- **prediktivní hodnota negativního testu** je pravděpodobnost, že osoba skutečně není nemocná, když test vyjde negativní, tj. $P(H^-|A^-)$. [2]

Tabulka (A.1) objasňuje, jak lze ze zjištěných dat spočítat hodnotu zmíněných ukazatelů správnosti (jedná se vlastně o vypočítání podmíněné pravděpodobnosti na základě daných jevů – výsledků testu, a skutečné ověřené přítomnosti nemoci):

výsledek testu	ověření přítomnosti nemoci / látky		celkem
	přítomna (H^+)	*nepřítomna (H^-)*	
pozitivní (A^+)	a	b	a + b
negativní (A^-)	c	d	c + d
celkem	a + c	b + d	a + b + c + d

Tab. A.1: Obecná tabulka zjištěných hodnot diagnostického testu.

Z tabulky (A.1) je zřejmé, že a je počet lidí s pozitivním výsledkem testu, u nichž je nemoc přítomna, b je počet lidí s negativním výsledkem, u nichž je nemoc nepřítomna, c je počet lidí s negativním výsledkem, u nichž je nemoc přítomna a d je počet lidí s pozitivním výsledkem, u nichž je nemoc nepřítomna.

Hodnoty ukazatelů v procentech se poté určují pomocí podmíněné pravděpodobnosti jako:

$$\begin{aligned} P(A^+|H^+) &= \frac{a}{a+c} \cdot 100, \\ P(A^-|H^-) &= \frac{d}{b+d} \cdot 100, \\ P(H^+|A^+) &= \frac{a}{a+b} \cdot 100, \\ P(H^-|A^-) &= \frac{d}{c+d} \cdot 100. \end{aligned}$$ [1]

V následujícím jednoduchém příkladu si ukážeme určení ukazatelů správnosti diagnostických testů.

Příklad A.3.1. Vypočítejte senzitivitu, specificitu a prediktivní hodnoty hypotetického testu na zjištění rakoviny na základně hodnot uvedených v tabulce (**??**). Jev A^+ značí, že náhodně vybraný pacient je pozitivní na rakovinu a opačný jev A^-, že osoba je negativní na rakovinu. Jev H^+ značí, že je rakovina u pacienta skutečně přítomna a opačný jev H^-, že je rakovina nepřítomna. Určete, zda je test přesný (jako přesný test zde budeme považovat test s hodnotami vyššími než 80 %).

Řešení:

Z tabulky (A.2) nejprve vypočítáme senzitivitu a specificitu testu:

$$P(A^+|H^+) = \frac{47}{52} \cdot 100 \doteq 90,4\,\%, \qquad P(A^-|H^-) = \frac{38}{42} \cdot 100 \doteq 90,5\,\%,$$

poté hodnotu pozitivního a negativního testu:

$$P(H^+|A^+) = \frac{47}{51} \cdot 100 \doteq 92,2\,\%, \qquad P(H^-|A^-) = \frac{38}{43} \cdot 100 \doteq 88,4\,\%.$$

výsledek testu	ověření přítomnosti rakoviny		celkem
	přítomna (H^+)	*nepřítomna (H^-)*	
pozitivní (A^+)	47	4	51
negativní (A^-)	5	38	43
celkem	52	42	94

Tab. A.2: Zjištěné hodnoty hypotetického testu na určení rakoviny.

Vidíme, že daný test je poměrně kvalitní a přesný, neboť se hodnoty ukazatelů pohybují nad $80\ \%$. Pravděpodobnost, že je pacient skutečně nemocný, když mu test vyjde pozitivní, je dokonce $92,2\ \%$.

Senzitivita a specificita jsou především populační ukazatele, protože se zjišťují na základě znalosti skutečné přítomnosti (resp. nepřítomnosti) onemocnění u určitého souboru lidí. Zajímají zejména lékaře, kteří je pak mohou použít pro porovnání diagnostické správnosti dvou různých testů. U konkrétního pacienta, který přijde do ordinace s výsledky testu, skutečnou přítomnost (resp. nepřítomnost) nemoci lékař nezná. [2]

Pro pacienty jsou zajímavější ukazatele prediktivní hodnoty, které vycházejí z určitého výsledku jejich testu, a které v případě, že jejich vlastní test vyšel pozitivně (resp. negativně), určují s jakou pravděpodobností danou nemoc skutečně mají (resp. nemají). [2]

Stejně tak jako v jiných statistických metodách, i u těchto testů zvyšuje jejich kvalitu dostatečná velikost experimentálního vzorku (tj. velikost souboru testovaných lidí). Pokud by byl tento soubor malý, poroste pravděpodobnost, že někteří specifičtí pacienti nebudou zachyceni a odhady specificity a senzitivity budou zkreslené. [2]

Mezní hodnoty senzitivity a specificity určující, zda je daný test kvalitní, resp. kvalitnější než jiné testy, které jsou aktuálně dostupné, se určují obtížně. Závisí totiž na míře poznání dané oblasti a na dosažitelné správnosti dostupných testů. V některé oblasti mohou být hodnoty nad 60 % vítězstvím, v jiné se diagnostika blíží v obou ukazatelích hodnotě 100 %. [2]

V dalším příkladu budeme porovnávat dva těhotenské testy. Těhotenské testy fungují na principu zjištění přítomnosti hormonu gonadotropinu v moči nebo v krevním séru. Gonadotropin je produkován buňkami vyvíjející se placenty a jeho koncentrace v moči či krvi v prvních týdnech po otěhotnění výrazně stoupá téměř každým dnem.

Příklad A.3.2. Dva hypotetické těhotenské testy – Mimicheck a Pregnus – na svém obalu tvrdí, že jsou nejlepším volně dostupným těhotenským testem. Úkolem je určit, který z nich je lepší (tj. spočítat a porovnat ukazatele správnosti).

K dispozici jsou výsledky obou těchto testů a výsledky ověření těhotenství, které byly zjišťovány na stejném souboru 200 žen (z toho 66 žen bylo lékařským vyšetřením ověřeno jako skutečně těhotných), zanesených v tabulkách (A.3):

výsledek testu Mimicheck	ověření skutečného těhotenství		celkem
	těhotná (H^+)	*není těhotná (H^-)*	
pozitivní (A^+)	63	4	67
negativní (A^-)	3	130	133
celkem	66	134	200

výsledek testu Pregnus	ověření skutečného těhotenství		celkem
	těhotná (H^+)	*není těhotná (H^-)*	
pozitivní (A^+)	60	2	62
negativní (A^-)	6	132	138
celkem	66	134	200

Tab. A.3: Zjištěné hodnoty pro Mimicheck a Pregnus.

Řešení:

Nejdříve vyhodnotíme Mimitest:

- senzitivita ... $P(A^+|H^+) = \dfrac{63}{66} \cdot 100 \doteq 95,5\ \%$
- specificita ... $P(A^-|H^-) = \dfrac{130}{134} \cdot 100 \doteq 97\ \%$
- prediktivní hodnota pozitivního testu ... $P(H^+|A^+) = \dfrac{63}{67} \cdot 100 \doteq 94\ \%$
- prediktivní hodnota negativního testu ... $P(H^-|A^-) = \dfrac{130}{133} \cdot 100 \doteq 97,7\ \%$

Poté vyhodnotíme Pregnus:

- senzitivita ... $P(A^+|H^+) = \dfrac{60}{66} \cdot 100 \doteq 90,9\ \%$
- specificita ... $P(A^-|H^-) = \dfrac{132}{134} \cdot 100 \doteq 98,5\ \%$
- prediktivní hodnota pozitivního testu ... $P(H^+|A^+) = \dfrac{60}{62} \cdot 100 \doteq 96,8\ \%$
- prediktivní hodnota negativního testu ... $P(H^-|A^-) = \dfrac{132}{138} \cdot 100 \doteq 95,7\ \%$

Porovnáním obou testů zjišťujeme, že oba testy vykazují vysoké pravděpodobnosti a jsou tedy oba poměrně kvalitní. Mimitest má o něco vyšší senzitivitu, to znamená, že s větší pravděpodobností rozpozná skutečně těhotné ženy. Pregnus má zase o $1,5$ % specificitu, což znamená, že je o něco citlivější na vyloučení žen, které nejsou těhotné. Navíc Mimitest vykazuje větší pravděpodobnost, že je žena skutečně těhotná, když test vyjde pozitivní, což je většinou hledisko, které nejvíce zajímá právě ženy provádějící těhotenské testy.

Poznámka: Ve skutečnosti příbalové letáky těhotenských testů uvádějí až 100 % *hodnoty senzitivity a specificity. V zadání předchozího příkladu byly hodnoty záměrně sníženy, aby bylo možné testy porovnat.*

A.3.2 Vliv prevalence na prediktivní hodnoty testu

Prediktivní hodnoty pozitivního a negativního testu mj. úzce souvisí s **prevalencí**[1] určité sledované nemoci (nebo obecně nějakého defektu) v dané populaci. V určitém časovém okamžiku, lze prevalenci vyjádřit jako procento pacientů se sledovanou nemocí určené ze všech osob v dané populaci. [2]

Jak moc mohou prediktivní hodnoty záviset na velikosti prevalence si ukážeme na řešeném příkladu se záměrně vybranými dvěma populacemi s extrémně rozdílnými počty nemocných jedinců.

Než se začneme zabývat příkladem, musíme si vyjádřit prediktivní hodnoty pomocí prevalence. Označme prevalenci jako $P(H^+)$, tj. pravděpodobnost, že určitá osoba je skutečně nemocná (v daný okamžik a v dané populaci) a platí $1-P(H^+) = P(H^-)$, kde $P(H^-)$ je pravděpodobnost, že určitá osoba není nemocná. Prediktivní hodnoty si nyní vyjádříme pomocí upraveného Bayesova vzorce (2.7) s úplnou pravděpodobností a pomocí hodnot senzitivity a specificity :

$$P(H^+|A^+) = \frac{P(A^+|H^+)P(H^+)}{P(A^+|H^+)P(H^+) + P(A^+|H^-)P(H^-)},$$

$$P(H^-|A^-) = \frac{P(A^-|H^-)P(H^-)}{P(A^-|H^-)P(H^-) + P(A^-|H^+)P(H^+)}.$$

Příklad A.3.3. Úkolem je určit a porovnat pozitivní a negativní prediktivní hodnoty hypotetického testu na zjištění HIV pozitivity, u kterého je výrobcem garantována senzitivita 98 %

[1]Prevalence je počet existujících nemocí či zdravotních problémů ve vybrané populaci k určitému datu.

a specificita $99\ \%$ pro populaci Lesotha, kde se k 1. 1. 2017 vyskytuje $25\ \%$ HIV pozitivních a poté pro populaci České republiky, kde se k 1. 1. 2017 vyskytuje pouze $0,03\ \%$ HIV pozitivních.

Řešení:

Nejprve budeme uvažovat populaci Lesotha s relativně vysokou prevalencí HIV pozitivity:

- $P(H^+) = 0,25$:

$$P(H^+|A^+) = \frac{0,98 \cdot 0,25}{0,98 \cdot 0,25 + (1-0,99) \cdot (1-0,25)} \doteq 0,970$$
$$P(H^-|A^-) = \frac{0,99 \cdot (1-0,25)}{0,99 \cdot (1-0,25) + (1-0,98) \cdot 0,25} \doteq 0,993$$

A poté populaci České republiky s relativně nízkou prevalencí HIV pozitivity:

- $P(H^+) = 0,0003$:

$$P(H^+|A^+) = \frac{0,98 \cdot 0,0003}{0,98 \cdot 0,0003 + (1-0,99) \cdot (1-0,0003)} \doteq 0,029$$
$$P(H^-|A^-) = \frac{0,99 \cdot (1-0,0003)}{0,99 \cdot (1-0,0003) + (1-0,98) \cdot 0,0003} \doteq 0,999.$$

Je patrné, že v populaci s relativně vysokou prevalencí HIV pozitivity má kvalitní test vysokou pozitivní i negativní prediktivní hodnotu, tj. osoby s pozitivním (resp. negativním) testem mají vysokou pravděpodobnost, že jsou skutečně HIV pozitivní (resp. negativní).

Ovšem v populaci s nízkou prevalencí HIV pozitivity, má kvalitní test sice vysokou negativní prediktivní hodnotu – na $99,9\ \%$ jsou osoby s negativním výsledkem testu opravdu HIV negativní, ale relativně nízkou pozitivní prediktivní hodnotu – pouze na $2,9\ \%$ jsou osoby s pozitivním testem skutečně HIV pozitivní.

Kvalita testu daná vysokou senzitivitou a specificitou je tedy velice relativní a hodně záleží na velikosti prevalence uvažované nemoci v této populaci.

A.4 Návrh pracovního listu

Téma: Podmíněná pravděpodobnost v lékařské diagnostice

Jméno a třída: Datum:

Příklad č. 1 Hypotetický drogový test D7 slouží pro kvalitativní stanovení drogových metabolitů v lidské moči. Jedna z testem detekovaných látek je THC (Tetrahydrocannabinol, což je hlavní psychoaktivní látka nacházející se v konopí, mj. pochází z drog marihuana a hašiš). Detekční mez uvedená výrobcem je pro THC-kannabinol, 11-nor-8-THC-9-COOH a 11-nor-9-THC-COOH je 50 ng/ml. Výrobce udává $92,8$ % specificitu a 95 % senzitivitu.

Vysvětli pomocí pravděpodobnosti, co znamená specificita a senzitivita na tomto konkrétním diagnostickém testu: ....

....

....

....

....

Na základě udaných hodnot senzitivity a specificity rozhodni, zda je možné test pokládat za kvalitní a přesný: ....

....

....

....

Příklad č. 2 Hypotetický samodiagnostický test potravinové intolerance AlergSeeker je velmi jednoduchý a rychlý. Testovací položka s nanesenými potravinovými extrakty během 40 minut ukáže, na kterou z potravin má člověk vytvořené protilátky (tj. je alergický). Stačí k tomu pouze malá kapička krve ze špičky prstu, která se rozpustí v roztoku. Tak lze identifikovat přítomnost protilátek na jednu či více potravin v závislosti na výskytu modře zbarvených bodů na reakční podložce. Mezi testované potraviny patří obiloviny, luštěniny, ořechy, maso a ryby, zelenina a ovoce, vejce, kravské mléko aj. V následující tabulce jsou uvedeny zjištěné hodnoty pro testování přítomnosti protilátek na laktózu v kravském mléce a na gluten (lepek) v obilovinách. Obě testování byla provedena na souboru 300 lidí.

výsledek testu (laktóza)	ověření skutečné intolerance		celkem
	přít. protilátek (H^+)	*nepř. protilátek (H^-)*	
pozitivní (A^+)	44	1	45
negativní (A^-)	1	254	255
celkem	45	255	300

výsledek testu (lepek)	ověření skutečné intolerance		celkem
	přít. protilátek (H^+)	*nepř. protilátek (H^-)*	
pozitivní (A^+)	19	0	19
negativní (A^-)	2	279	281
celkem	21	279	300

Vypočítej všechny ukazatele správnosti u obou testování a porovnej prediktivní hodnotu pozitivního testu (pravděpodobnost, že osoba je skutečně alergická, když test vyjde pozitivní) při zjišťování přítomnosti protilátek na kravské mléko a na lepek.

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

Příklad č. 3 V následující tabulce jsou uvedeny čtyři neinvazivní zátěžové kardiotesty, což jsou testy činnosti srdce. Neinvazivní znamená, že se při jejich provádění nezasahuje dovnitř organismu. **Doplň chybějící data v tabulce a na základě senzitivity a specificity porovnej kvalitu těchto testů.**

	druh testu	počet pacientů	výsledek testu	ověření přítomnosti nemoci: *přítomna (H^+)*	ověření přítomnosti nemoci: *nepřítomna (H^-)*	senzitivita (%)	specificita (%)
1.	**zátěžová EKG**	**11671**	*pozitivní (A^+)*		2345		**71,2**
			negativní (A^-)	1047	5790		
2.	**zátěžový SPECT**	**3343**	*pozitivní (A^+)*	1024	641		
			negativní (A^-)	433	1245		
3.	**zátěžová echokardiografie**		*pozitivní (A^+)*	360			**74,2**
			negativní (A^-)	59	224		
4.	**zátěžová magnetická rezonance**		*pozitivní (A^+)*	294		**90,2**	**85,2**
			negativní (A^-)		271		

. . . .

. . . .

. . . .

A.5 Autorské řešení pracovního listu

Příklad č. 1 Senzitivita drogového testu $92,8\ \%$ znamená pravděpodobnost $0,928$, že tento test bude pozitivní, když má osoba provádějící test skutečně v těle takové množství THC, které je test schopný detekovat. Specificita drogového testu $95\ \%$ znamená pravděpodobnost $0,95$, že tento test bude negativní, když nemá osoba provádějící test v těle žádné množství THC nebo když má v těle nižší množství THC, než je test schopný detekovat.

Na základě hodnot senzitivity a specificity, které jsou vyšší než $90\ \%$, lze usoudit, že je test relativně kvalitní a přesný.

Příklad č. 2 Intolerance laktózy: $P(A^+|H^+) \doteq 97,8$ %, $P(A^-|H^-) \doteq 99,6$ %, $P(H^+|A^+) \doteq 97,8$ %, $P(H^-|A^-) \doteq 99,6$ %.
Intolerance lepku: $P(A^+|H^+) \doteq 90,5$ %, $P(A^-|H^-) = 100$ %, $P(H^+|A^+) = 100$ %, $P(H^-|A^-) \doteq 99,3$ %.

Test vykazuje u zjišťování intolerance laktózy i lepku vysoké hodnoty. Prediktivní hodnota pozitivního testu je vyšší u zjišťování alergie na lepek (100 %).

Příklad č. 3 Hodnoty lze snadno dopočítat z definic senzitivity a specificity.

	druh testu	počet pacientů	zjištěné hodnoty: výsledek testu	ověření přítomnosti nemoci: přítomna (H^+)	ověření přítomnosti nemoci: nepřítomna (H^-)	senzitivita (%)	specificita (%)
1.	zátěžová EKG	11671	pozitivní (A^+)	2489	2345	70,4	71,2
			negativní (A^-)	1047	5790		
2.	zátěžový SPECT	3343	pozitivní (A^+)	1024	641	70,3	66
			negativní (A^-)	433	1245		
3.	zátěžová echokardiografie	721	pozitivní (A^+)	360	78	86	74,2
			negativní (A^-)	59	224		
4.	zátěžová magnetická rezonance	644	pozitivní (A^+)	294	47	90,2	85,2
			negativní (A^-)	32	271		

Jako nejméně kvalitní zátěžový kardiotest se jeví SPECT, naopak jako nejvíce kvalitní se jeví magnetická rezonance.

A.6 Závěr

V úvodu čtenář zjistil především to, že podmíněná pravděpodobnost má široké uplatnění v lékařské diagnostice a že aplikace podmíněné pravděpodobnosti při ověřování diagnostických testů lze vhodně použít jak v hodinách středoškolské matematiky, tak i biologie.

V druhé části článku se čtenář stručně seznámil s teorií podmíněné pravděpodobnosti a s Bayesovým vzorcem.

V třetí části bylo čtenáři nejprve vysvětleno, co je to diagnostický test, jaké jsou jeho ukazatele správnosti a na řešených příkladech mu byl představen princip vyhodnocování těchto testů. Poté se dozvěděl, že senzitivita a specificita jsou charakteristiky samotného diagnostického testu, ale prediktivní hodnoty jsou velmi ovlivněny prevalencí – tj. tím, jak často se nemoc vyskytuje v populaci v určitém okamžiku, což bylo demonstrováno na posledním příkladu této části.

Poslední část s návrhem pracovního listu a jeho řešením nám nabídla čtenáři další úlohy pro procvičení dané problematiky.

K napsání článku byl použit sázecí systém LaTeX.

Literatura

[1] ZVÁROVÁ J. *Základy statistiky pro biomedicínské obory.* Nakladatelství Karolinum, Praha, 2004.

[2] Senzitivita, specificita a prediktivní hodnoty. *Matematická biologie: E-learningová učebnice* [online]. [cit. 2017-03-15].

Dostupné z: Dostupné z: http://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat–biostatistika-pro-matematickou-biologii–vztah-pravdepodobnosti-statistiky-a-biostatistiky–senzitivita-specificita-a-prediktivni-hodnoty

[3] Hátle, J., Kahounová, J.: *Úvod do teorie pravděpodobnosti*. SNTL/ALFA, Praha, 1987.

[4] Dupač V., Hájek J. *Pravděpodobnost ve vědě a technice*. 1. vydání, Academia, 1962

[5] Podmíněná pravděpodobnost a Bayesův vzorec. *Matematická biologie: E-learningová učebnice* [online]. [cit. 2017-03-15].

Dostupné z: http://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat–biostatistika-pro-matematickou-biologii–vztah-pravdepodobnosti-statistiky-a-biostatistiky–podminena-pravdepodobnost-a-bayesuv-vzorec

[6] Kapitola 3 Diagnostika. *Medicína založená na důkazu* [online]. [cit. 2017-03-26].

Dostupné z: Dostupné z: http://www.khshk.cz/e-learning/kurs1b/kapitola_3__diagnostika.html

Bc. Lucie Studená

UMB, PřF JU v Českých Budějovicích

Branišovská 1760, 370 05 České Budějovice

e-mail: lucine.studena@seznam.cz