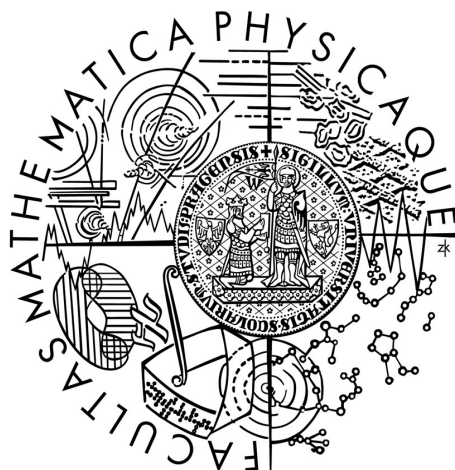


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Martina Selementová

STOCHASTICKÉ MODELY PRO GENETICKOU ANALÝZU

Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: Mgr. Michal Kulich, PhD.

Studijní program: Matematika
Studijní obor: Finanční matematika

2006

Na tomto místě chci poděkovat vedoucímu své bakalářské práce, panu Mgr. Michalovi Kulichovi, PhD., za zapůjčení literatury a dále za všechny cenné rady a připomínky, které mi poskytl.

Prohlašuji, že jsem svou bakalářskou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze 25. května 2006

Martina Selementová

Obsah

1 Úvod	4
1. 1 Úvod do obecné genetiky	4
2 Náhodné křížení populací	6
2. 1 Koncept četnosti genů	6
2. 2 Náhodné křížení	6
2. 3 Pohlavně vázané geny	8
2. 4 Tetraploidní organismy	10
2. 5 Autosterilizační geny	11
2. 6 Dva lokusy, každý se dvěma alelami	13
2. 7 Obecné chování dvou provázaných lokusů	15
3 Základní problémy výběru	18
3. 1 Výběr založený na genotypu	18
3. 2 Výběr omezený na pohlaví	20
3. 3 Selektce pohlavně vázaných genů	21
3. 4 Jiné případy	23
4 Závěr	25
Použitá literatura a software	26

Název práce: Stochastické modely pro genetickou analýzu

Autor: Martina Selementová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Michal Kulich, PhD.

e-mail vedoucího: kulich@karlin.mff.cuni.cz

Abstrakt: V předložené práci studujeme elementární stochastické metody pro modelování genetických populací. Nejprve uvádíme základní pojmy obecné genetiky pro snadnější orientaci v textu. V následujících kapitolách se zabýváme dynamikou vývoje populace a základními problémy výběru v populaci. Hlavním tématem je náhodné křížení v populaci, které dále aplikujeme na pohlavně vázané geny, tetraploidy, autosterilizační geny a na případ dvou lokusů. Ve druhé části se zabýváme problémy výběru: zaměřili jsme se především na výběr založený na genotypu aplikovaný na pohlaví a pohlavně vázané geny a dále na familiární výběr a konkrétní příklad Rh faktoru.

Klíčová slova: náhodné křížení, výběr

Title: Stochastic Models for Genetic Analysis

Author: Martina Selementová

Department: Department of Probability and Mathematical Statistics

Supervisor: Mgr. Michal Kulich, PhD.

Supervisor's e-mail address: kulich@karlin.mff.cuni.cz

Abstract: In the presented work we study elementary stochastic methods for modeling genetic populations. First we show some basic notion of general genetics for easier orientation in the text. In the following chapters we deal with the progress of population dynamics and the elementary problems of selection. Our main topic is the random mating within population which we apply to sex-linked genes, autotetraploids, self-sterility genes and the case of two loci. In the second part we deal with problems of selection: we target selection based on genotype applied to sex-limited selection and sex-linked genes and eventually familial selection and particular example of Rhesus factor.

Keywords: random mating, selection

1 Úvod

Cílem této práce je přiblížit základní chování genetických populací pomocí jednoduchých stochastických modelů. Zajímá nás, jak se daná populace organismů bude za určitých podmínek vyvíjet z generace na generaci. Mohou to být podmínky uvnitř dané populace (konkrétní vlastnosti jedinců), anebo podmínky okolního prostředí (např. zásah chovatele). Zde je zásadní porozumět dědičnosti a teorii evoluce. Chceme tedy ukázat, že dědičnost v populacích je možné popsat jednotným mechanismem – že pravidla dědičnosti jsou neměnná. Zaměříme se tedy na chování uvnitř jednotlivých populací, kde populací rozumíme izolovanou skupinu jedinců stejného druhu. Toto lze nejnadhěji prezentovat pomocí *náhodného křížení*. Co se týče vývoje v populaci je také zajímavý *problém výběru*. Tedy situace, kdy jsou někteří jedinci preferováni (jsou vyloučeni jedinci s nějakým konkrétním znakem).

Hlavní myšlenka této práce je převzata z knihy Osacara Kempthorna *An Introduction to genetic Statistics* (Iowa State University Press, Ames, 1969).

Nejprve zde zavedeme koncept četnosti genů a dále se seznámíme s náhodným křížením. Náhodné křížení pak v následujících kapitolách aplikujeme na pohlavně vázané geny, tetraploidní organismy, autosterilizační geny a na případ dvou lokusů. Následně se zaměříme na základní problémy výběru – na výběr založený na genotypu, a to omezený na pohlaví a opět na pohlavně vázané geny. V poslední části ukážeme familiární výběr a problém Rh faktoru.

Veškerá data použitá v prezentovaných příkladech jsou smyšlená, a nejedná se tedy o žádné vlastnosti skutečných populací.

1. 1 Úvod do obecné genetiky

Genetika je biologická věda zabývající se dědičností a proměnlivostí organismů. Díky dědičnosti se potomci podobají svým rodičům více než jakýmkoli jiným jedincům. Proměnlivost (variabilita) populace jedinců je dána jak vnějšími podmínkami, tak rozdíly v dědičných vlastnostech jedince. A právě dědičnost a rozmanitost souvisí se schopnostmi rozmnožování (reprodukce), tj. schopností vytvářet nové jedince.

Rozlišujeme dva typy rozmnožování

1. nepohlavní. Jedinec vzniká přímo z mateřské buňky, části tkáně či orgánu. Generace vzniklá nepohlavním rozmnožováním je nazývána klony.
2. pohlavní. Nový jedinec zde vzniká splynutím dvou pohlavních haploidních buněk (gamet) rodičů. Generace vzniklé pohlavním rozmnožováním označujeme jako potomstvo.

Pro všechno naše následující užití předpokládáme rozmnožování pohlavní.

Gameta (pohlavní buňka) má jen jednu sadu nepárových chromozomů – je haploidní – takže po splynutí dvou gamet vzniká klasicky diploidní organismus (s jednou sadou párových chromozomů). Gametickým polem je označen soubor veškerých gamet vyprodukovaných jedincem či populací.

Veškeré znaky organismu jsou řízeny geny. Gen je jednotka informace a soubor veškerých genů jednoho jedince označujeme jako jeho genotyp. My zde budeme označovat jako genotyp pouze sestavu těch genů, které budeme sledovat. Projev genů

pozorujeme prostřednictvím fenotypu, tj. projevujících se vlastností. Termín genotypové pole označuje soubor všech genotypů v populaci a jejich zastoupení.

Gen se může v populaci vyskytnout v různých formách – alelách. V tělesné buňce diploidního organismu je každý gen zastoupen dvěma alelami, z nichž jedna pochází ze samčího a druhá ze samičího organismu. Gamety nesou tedy vždy po jedné alele. Jsou-li obě alely stejného typu, označujeme jejich sestavu jako homozygotní, jsou-li rozdílné tak jako heterozygotní. Při křížení rodičů lišících se v jenom znaku (např. barvě květu) má celé heterozygotní potomstvo vzniklé tímto zkřížením vzhled jednoho z rodičů. Znak, který se v potomstvu objevil, je označen jako dominantní a ten neprojevený jako recesivní. V textu se dále budeme držet značení dominantních znaků verzálkami a recesivních znaků minuskami.

Geny jsou uloženy na chromozomech. Chromozomy jsou vláknité útvary nacházející se v jádře každé eukaryotické buňky, obsahující kyselinu deoxyribonukleovou. Místo, ve kterém se jednotlivé alely vážou na chromozom, je nazýváno lokus.

Nositelkami genetické informace jsou nukleové kyseliny. U většiny organismů je touto nositelkou deoxyribonukleová kyselina (DNA), u RNA-virů kyselina ribonukleová (RNA). Nukleové kyseliny jsou polymerní látky, jejichž základními složkami jsou nukleotidy. Nukleotid vznikne spojením organické báze, kyseliny trihydrogenfosforečnaté a pětiuhlíkatého cukru. Jednotlivé nukleotidy jsou spojeny fosfáty do polynukleotidového řetězce.

Zakladatelem genetiky je Johann Gregor Mendel (1822 – 1884) narozený v Hynčicích na Moravě. Byl zvolen opatem brněnského kláštera, kde prováděl pokusy s křížením rostlin. Podařilo se mu experimentálně dokázat, že dědičnost lze objasnit jednotným mechanismem. Výsledek jeho pokusů byl však uznán až v roce 1900, kdy nezávisle na sobě a téměř současně objevili Mendelem formulované zákony dědičnosti De Vries, Correns a Tschermak.

2 Náhodné křížení populací

Zde se budeme snažit popsat použití teorie pravděpodobnosti na jednoduchý problém populační dynamiky. To jest, jak se populace za určitých podmínek mění z generace na generaci. Konkrétně nás bude například zajímat, jak se populace, počínající zkřížením $AB \times AB$, bude vyvíjet v následujících generacích.

V křížení, které můžeme označit jako nultou generaci, je podíl jednotlivců ve fázi páření 100%. Můžeme se pak ptát, jaké budou podíly různých pářících se jedinců v desáté či sté generaci. V odpovědi na tuto otázku je zásadní porozumět evoluční teorii a dědičnosti.

2.1 Koncept četnosti genů

Předpokládejme, že máme dvě alely na lokusu, tři rozlišitelné genotypy a máme dány počty zastoupení jednotlivých genotypů v populaci.

Genotyp	Počet	Podíl
AA	N_2	$P = N_2/N$
Aa	N_1	$2Q = N_1/N$
aa	N_0	$R = N_0/N$
Celkem	N	

Jako genotypové pole je označen soubor veškerých genotypů v populaci a jejich zastoupení, takže v tomto případě genotypové pole zapíšeme jako $PAA + 2QAa + Raa$.

Z celkového počtu $2N$ genů v populaci je $2N_2 + N_1$ typu A a $N_1 + 2N_0$ typu a . Tedy četnost genů A je $P_A = \frac{2N_2 + N_1}{2N} = P + Q$ a podobně $P_a = Q + R$. Tyto četnosti pak vyjadřují počet jednotlivých gamet A a a vyprodukovaných populací. Gametické pole (soubor veškerých gamet v populaci) zapíšeme jako $(P + Q)A + (Q + R)a$.

Předpokládejme, že velikost populace je n s pozorovanými četnostmi $p_s, 2q_s, r_s$ genotypů AA, Aa a a . P_A pak můžeme odhadnout jako

$$\hat{P}_A = p_s + q_s .$$

2.2 Náhodné křížení

Nyní můžeme uplatnit zákonitosti pravděpodobnosti ve studiu vývoje populací co se týče četnosti genů, četnosti genotypů atp.

Nejprve uveďme větu, kterou Kempthorne považuje jako základní a která proces výrazně zkrátí:

Genotypové pole populace vzniklé náhodným křížením je čtvercem gametického pole v prvotní populaci, kde se každý z rodičů podílí stejným dílem.

Důkaz: Pokud podíl počtu jednotlivců typu A_1 je P_1 , A_2 je P_2 atd, pak $\sum P_i A_i$ je nazývána genotypovým polem populace. My zde potřebujeme jak genotypové tak gametické pole.

Předpokládejme, že v prvotní populaci g máme genotypy G_1, G_2, \dots, G_g v poměru P_1, P_2, \dots, P_g a jedinec i -tého genotypu produkuje gamety $\alpha_1, \alpha_2, \dots, \alpha_s$ v poměru $p_{i1}, p_{i2}, \dots, p_{is}$. Potom pravděpodobnost zkřížení $G_i \times G_j$ bude $P_i P_j$ a pravděpodobnost, že při tomto křížení splynou gamety α_m a α_n je $p_{im} p_{jn}$ (dejme tomu, že α_m označuje gametu A a α_n a , pak se s pravděpodobností $p_{im} p_{jn}$ setkají a vznikne jedinec typu $\alpha_m \alpha_n$ – tedy typu Aa). Odtud je celkové genotypové pole výsledné populace

$$\sum_{ij} P_i P_j \sum_{mn} p_{im} p_{jn} \alpha_m \alpha_n,$$

kde $P_i P_j p_{im} p_{jn}$ je příspěvek jedince s genotypem $\alpha_m \alpha_n$ z křížení $G_i \times G_j$.

$$\sum_{ij} P_i P_j \sum_{mn} p_{im} p_{jn} \alpha_m \alpha_n = \sum_{mn} \left(\sum_i P_i p_{im} \alpha_m \right) \left(\sum_j P_j p_{jn} \alpha_n \right) = \left(\sum_m \sum_i P_i p_{im} \alpha_m \right)^2,$$

kde $\sum_m \sum_i P_i p_{im} \alpha_m$ je gametické pole prvotní populace, jelikož podíl produkovaných gamet α_m je $\sum_i P_i p_{im}$. Každý jedinec genotypu G_i totiž produkuje p_{im} gamet α_m . \square

Termín náhodné křížení zahrnuje dvě podmínky. První z nich je ta, že ke vzájemnému křížení dochází uvnitř *jedné* populace. Druhou je předpoklad, že z každého zkřížení vzejde *stejný počet* potomků. Různý počet potomků je problém, ve kterém je rozhodující výběr. A samozřejmě, jedinci se kříží zcela náhodně – neexistuje zde žádná preference. Při použití této věty je tedy třeba dostát všem předpokladům, na kterých je založena.

Je zde tedy snadno vidět, že genotypové pole potomků je rovno součinu gametického pole samce (otce) a gametického pole samice (matky). Rovnováhy v populaci je dosaženo tehdy, kdy se četnosti jednotlivých genů při náhodném křížení nemění.

Příklad: Uvažujme zkřížení následujících dvou populací:

$$\begin{aligned} \text{I} & 0,25AA + 0,5Aa + 0,25aa, \\ \text{II} & 0,16AA + 0,48Aa + 0,36aa. \end{aligned}$$

Teoreticky uvažujme, že A představuje bílé zbarvení květu a a růžové. U heterozygotních jedinců převládá znak A . A je tedy dominantní zatímco a recesivní znak (i nadále budeme tuto vlastnost rozlišovat stejným způsobem – tj. verzálkou dominantní znak a minuskou recesivní). V populaci I je pak 75 % rostlin bělokvětvých a 25 % růžovokvětvých, v populaci II 64 % bělokvětvých a 36 % růžovokvětvých.

Samce vezmeme z I a samici z II.

Gametická pole jsou tedy následující:

$$\begin{aligned} \text{I} & 0,5A + 0,5a, \\ \text{II} & 0,4A + 0,6a. \end{aligned}$$

Genotypové pole výsledné populace je pak:

$$(0,5A + 0,5a)(0,4A + 0,6a) = 0,2AA + 0,5Aa + 0,3aa.$$

Tudíž výslednou populaci bude představovat 70 % rostlin s bílými a 30 % s růžovými květy.

Pokud bychom obě populace podrobili náhodnému křížení (tzn., že křížení mezi populacemi by bylo zabráněno – nastávalo by pouze uvnitř jednotlivých populací),

jejich výsledná genotypová pole by byla následující:

$$\begin{aligned} \text{I} \quad & (0,5A + 0,5a)^2 = 0,25AA + 0,5Aa + 0,25aa, \\ \text{II} \quad & (0,4A + 0,6a)^2 = 0,16AA + 0,48Aa + 0,36aa, \end{aligned}$$

z čehož je patrné, že rovnováhy (nezměněných četností jednotlivých genů) je dosaženo ihned v první generaci náhodného křížení. Četnosti bílých a růžových květu se tedy z generace na generaci nemění.

2.3 Pohlavně vázané geny

Pohlavně vázaný gen je takový gen, který se projeví v závislosti na pohlaví jedince – například pouze u samců. V případě pohlavně vázaných genů rodiče potomkovi nepřispívají rovnoměrně. Samičí pohlaví je homogametní – samice mají dva X chromozomy, zatímco samci jsou heterogametní – mají jeden X a jeden Y chromozom. Obecná struktura přenosu na potomka je ta, že samec předá X chromozom své dceři a Y chromozom synovi. Samice předává náhodně svůj pohlavní chromozom bez ohledu na pohlaví potomka. Pohlavně vázaný gen je ten gen, který nosí chromozom X.

Syn tedy získává svůj X chromozom od své matky. Je tedy zřejmé, že jeho genotypové pole je dle náhodného křížení gametickým polem jeho matky. Zatímco dcery dědí stejným dílem od matky tak od otce – jejich genotypové pole je součinem gametických polí otce a matky. Pro ilustraci začneme lokusem se dvěma alelami v populaci skládající se z:

Samice	AA	Aa	aa
Četnost	r_0	$2s_0$	t_0
kde $r_0 + 2s_0 + t_0 = 1$			
Samci	A	a	
Četnost	p_0	q_0	
kde $p_0 + q_0 = 1$			

Předpokládejme, že tato data jsou četnostmi nulté generace. Chceme teď spočítat četnosti v generacích 1, 2, ... dle náhodného křížení.

Gametické pole samic je $(r_0 + s_0)A + (s_0 + t_0)a$ a samců $p_0A + q_0a$.

Tedy genotypové pole samic v generaci 1 je

$$\begin{aligned} & ((r_0 + s_0)A + (s_0 + t_0)a)(p_0A + q_0a) \\ & = (r_0 + s_0)p_0AA + (p_0(s_0 + t_0) + q_0(r_0 + s_0))Aa + q_0(s_0 + t_0)aa = r_1AA + 2s_1Aa + t_1aa \end{aligned}$$

a genotypové pole samců v generaci 1 je stejné jako samičí gametické pole v generaci 0, tedy

$$(r_0 + s_0)A + (s_0 + t_0)a = p_1A + q_1a.$$

Tímto rekurentním způsobem je možné vyjádřit i četnosti v generacích vyšších:

$$\begin{aligned} r_{n+1} &= p_n(r_n + s_n), & 2s_{n+1} &= p_n(s_n + t_n) + q_n(r_n + s_n), & t_{n+1} &= q_n(s_n + t_n), \\ p_{n+1} &= r_n + s_n, & q_{n+1} &= s_n + t_n, \end{aligned}$$

$$\text{tedy:} \quad r_{n+1} = p_n p_{n+1}, \quad 2s_{n+1} = p_n q_{n+1} + q_n p_{n+1}, \quad t_{n+1} = q_n q_{n+1}.$$

Příklad: Uvedme tabulku (*Tabulka 1*) vývoje populace při náhodném křížení s prvotními genotypovými četnostmi

Samci	$0,4A + 0,6a$	$p = 0,4$	$q = 0,6$	
Samice	$0,3AA + 0,5Aa + 0,2aa$	$r = 0,3$	$2s = 0,5$	$t = 0,2$

Tabulka 1

<i>Generace</i>	<i>r</i>	<i>2s</i>	<i>t</i>	<i>p</i>	<i>q</i>
0	0,3	0,5	0,2	0,4	0,6
1	0,22	0,51	0,27	0,55	0,45
2	0,26125	0,5025	0,23625	0,475	0,525
3	0,243438	0,500625	0,255938	0,5125	0,4875
4	0,253047	0,500156	0,246797	0,49375	0,50625
5	0,248418	0,500039	0,251543	0,503125	0,496875
10	0,250049	0,500000	0,249951	0,499902	0,500098
16	0,250001	0,500000	0,249999	0,499998	0,500002
17	0,250000	0,500000	0,250000	0,500001	0,499999
18	0,250000	0,500000	0,250000	0,500000	0,500000
19	0,250000	0,500000	0,250000	0,500000	0,500000

Z výsledků je patrné, že, při zadaných počátečních četnostech, dominantních samic ubývá a recesivních přibývá – u samců je tomu naopak. Pouze četnost heterozygotních samic se mezi populacemi relativně nemění. Populace se k bodu své rovnováhy přibližuje pomalu. V 18. generaci náhodného křížení bychom tento stav mohli považovat za rovnovážný s přihlédnutím k tomu, že jednotlivé rozdíly v genotypových četnostech v populaci v dalších generacích jsou téměř nezatelné (v uvedené tabulce jsou hodnoty zaokrouhlovány na šest desetinných míst).

Tato úprava může být ihned zobecněna na případ m alel ($m > 2$). Alely označíme A_i , kde i nabývá hodnot 1, 2, ..., m . Pak je samičí genotypové pole v generaci 0 rovno $\sum_{ij} P_{ij}^{(0)} A_i A_j$, kde $P_{ij}^{(0)}$ je četnost jednotlivých genotypů $A_i A_j$ v nulté generaci a $P_{ij}^{(0)} = P_{ji}^{(0)}$.

Samičí gametické pole je tedy $\sum_i (\sum_j P_{ij}^{(0)}) A_i = \sum_i P_i^{(0)} A_i$, zatímco samčí genotypové a gametické pole je identické: $\sum_i M_i^{(0)} A_i$.

Z těchto vztahů dále vyplývá $M_i^{(n+1)} = P_i^{(n)}$,

$$2 P_{ij}^{(n+1)} = P_i^{(n)} M_j^{(n)} + P_j^{(n)} M_i^{(n)} = P_i^{(n)} P_j^{(n-1)} + P_j^{(n)} P_i^{(n-1)}.$$

Použitím $2 \sum_j P_{ij}^{(n+1)} = 2 M_i^{(n+2)}$ dostáváme $2 M_i^{(n+2)} = M_i^{(n+1)} + M_i^{(n)}$, což je stejný vztah jako pro případ dvou alel.

V lidské populaci bylo zatím objeveno asi 50 těchto pohlavně vázaných genů. Patří mezi ně například alela podmiňující barvoslepost, hemofilii nebo absenci potních

žlázek. Vzhledem k tomu, že tyto znaky jsou vázány na nehomologických¹ částech chromozomu X, projevují se výrazným rozdílem v četnosti mezi pohlavími. Tyto znaky se tak mnohem častěji projevují u jedinců XY než u jedinců XX.

Například hemofilie se u mužů vyskytuje přibližně s frekvencí 1 : 10 000 a u žen je velice vzácná – s frekvencí přibližně 1 : 100 000 000 [3]. Je to dáno tím, že recesivní alela *h*, která hemofilii způsobuje, se projeví vždy. Muži tedy díky uložení alely *h* nikdy nemohou mít i alelu dominantní +, která by potlačila její projev. Naopak i žena, která nemocí netrpí, může být její heterozygotní přenašečkou +*h*. Pokud je tedy otec hemofilikem, jeho děti jsou zdravé, ale dcery přenašečkami. Je-li matka přenašečkou, pak její synové budou s pravděpodobností 50 % hemofilici a dcery v poměru 1 : 1 genotypu ++ a +*h*, tedy opět s pravděpodobností 50 % budou přenašečkami.

2. 4 Tetraploidní organismy

Působením alkaloidu kolchicinu či jinými způsoby lze u rostlin zdvojnásobit počet chromozomů (kolchicin naruší proces dělení buněčných jader tak, že chromozomy se ve svých párech nemohou rozejít [2]). Jedinci se zdvojnásobeným počtem chromozomů ve svých tělových buňkách jsou tetraploidní. U některých rostlin – např. jetel či žito – jsou tetraploidní jedinci větší a statnější než diploidní. V rostlinné výrobě je však takovéto šlechtění používáno jen v omezené míře, jelikož původní představa, že zdvojení chromozomů bude mít univerzálně za následek zvětšení buněk organismů, a tím i zvětšení některých jejich částí – např. semen – se ukázala mylnou.

Tetraploid má tedy zdvojenou sadu chromozomů – má čtyři geny v lokusu (řekněme *a*, *b*, *c* a *d*). Rodič předává svému potomkovi náhodně jeden ze šesti možných párů genů (*ab*, *ac*, *ad*, *bc*, *bd*, *cd*). Tedy každý gen je předán s pravděpodobností jedné poloviny. Při dvou alelách se vyskytuje pět genotypů produkujících následující gamety:

<i>Genotyp</i>	<i>Gamety</i>
$AAAA = A^4$	AA
$AAAa = A^3a$	$\frac{1}{2} AA + \frac{1}{2} Aa$
$AAaa = A^2a^2$	$\frac{1}{6} AA + \frac{4}{6} Aa + \frac{1}{6} aa$
$Aaaa = Aa^3$	$\frac{1}{2} Aa + \frac{1}{2} aa$
$aaaa = a^4$	aa

Uvažujme prvotní náhodně se křížící populaci, jejíž gametické pole je $x_0AA + 2y_0Aa + z_0aa$, takže populace v generaci 1 je $(x_0AA + 2y_0Aa + z_0aa)^2$.

Gametické pole generace 1 můžeme tedy vyjádřit jako $x_1AA + 2y_1Aa + z_1aa$, kde

$$x_1 = x_0^2 + 2x_0y_0 + \frac{2}{3} y_0^2 + \frac{1}{3} x_0z_0, \quad y_1 = x_0z_0 + \frac{4}{3} y_0^2 + \frac{2}{3} x_0z_0 + y_0z_0,$$

1 Pokud bychom vedle sebe srovnávali chromozom X a chromozom Y, bylo by patrné, že v jedné části se shodují, zatímco ve druhé se odlišují – mají různý počet vazbových skupin. A právě tyto shodné části jsou označeny jako homologické části chromozomů a ty rozdílené jako nehomologické. [2]

$$z_1 = \frac{2}{3} y_0^2 + \frac{1}{3} x_0 z_0 + 2y_0 z_0 + z_0^2.$$

Zde je snadné ověřit, že frekvence genů se nemění. Necht'

$$x_0 + y_0 = p = x_1 + y_1, \quad y_0 + z_0 = q = y_2 + z_1.$$

Nyní uvažujme

$$x_1 - p^2 = x_0^2 + 2x_0 y_0 + \frac{2}{3} y_0^2 + \frac{1}{3} x_0 z_0 - (x_0 + y_0)^2 = \frac{1}{3} (x_0 z_0 - y_0^2),$$

podobně $y_1 - pq = -\frac{1}{3} (x_0 z_0 - y_0^2), \quad z_1 - q^2 = \frac{1}{3} (x_0 z_0 - y_0^2).$

Hodnotu $(x_0 z_0 - y_0^2)$ označme Δ s indexem označujícím danou generaci, tedy $\Delta_i = (x_i z_i - y_i^2)$ pro i -tou generaci. Máme

$$x_1 = p^2 + \frac{1}{3} \Delta_0, \quad y_1 = pq - \frac{1}{3} \Delta_0, \quad z_1 = q^2 + \frac{1}{3} \Delta_0.$$

A tedy $\Delta_1 = (x_1 z_1 - y_1^2) = \frac{1}{3} \Delta_0, \quad \Delta_n = \left(\frac{1}{3}\right)^n \Delta_0.$

Tudíž četnosti jednotlivých genotypů v n -té generaci jsou $x_n = p^2 + \left(\frac{1}{3}\right)^n \Delta_0,$
 $y_n = pq - \left(\frac{1}{3}\right)^n \Delta_0, \quad z_n = q^2 + \left(\frac{1}{3}\right)^n \Delta_0.$

Genotypové pole generace $n + 1$ je rovno

$$\left\{ \left[p^2 + \left(\frac{1}{3}\right)^n \Delta_0 \right] AA + 2 \left[pq + \left(\frac{1}{3}\right)^n \Delta_0 \right] Aa + \left[q^2 + \left(\frac{1}{3}\right)^n \Delta_0 \right] aa \right\}^2.$$

Pokud použijeme limitní přechod, kde $n \rightarrow \infty$, genotypové pole bude

$$(p^2 AA + 2pq Aa + q^2 aa)^2 \text{ čili } (pA + qa)^4.$$

Rovnosti je zde dosaženo pouze asymptoticky, pokud však $\Delta_0 = 0$, rovnováha dle náhodného křížení nastává ihned.

Toto odvození pro dvě alely může být rozšířeno pro libovolný počet alel.

Příklad: Uvažujme prvotní populaci skládající se pouze z jedinců typu $AAaa$. Uvažujeme náhodné křížení a zajímá nás četnost jedinců $AAAA$ v následující generaci.

Gametické pole prvotní generace je tedy $\frac{1}{6} AA + \frac{4}{6} Aa + \frac{1}{6} aa$, tudíž genotypové pole výsledné populace je následující

$$\left(\frac{1}{6} AA + \frac{4}{6} Aa + \frac{1}{6} aa \right)^2 = \frac{1}{36} A^4 + \frac{8}{36} A^3 a + \frac{18}{36} A^2 a^2 + \frac{8}{36} A a^3 + \frac{1}{36} a^4.$$

Z toho je zřejmé, že četnost jedinců $AAAA$ je $\frac{1}{36}$.

2.5 Autosterilizační geny

Tento případ je zajímavý tím, že pravděpodobnostní argument je jednoduchý v případě tří alel. Necht' máme na lokusu alely S_1 , S_2 a S_3 . Pylová zrna obsahující alelu S_1 pak nemohou splnit svou reprodukční funkci na rostlinách, které také vlastní právě tuto alelu – tento gen označujeme jako autosterilizační. Předpokládejme, že všechny tři geny S_1 , S_2 a S_3 jsou autosterilizační, tudíž se v této populaci nevyskytují žádná homozygota. Rostlina s genotypem S_1S_2 může být opylena pouze S_3 pylem a vytváří S_1S_3 a S_2S_3 jedince ve stejném počtu.

Autosterilita je typická pro mnoho druhů ovocných stromů s jádrovými plody – např. jablň, hrušeň či kdoulevec [5]. Proto je třeba v těchto ovocných sadech vysadit i jinou odrůdu, která pro ně bude dobrým opylovačem.

Necht' genotypové pole prvotní populace je $x_0S_1S_2 + y_0S_1S_3 + z_0S_2S_3$.

Výsledné genotypové pole následující populace pak označíme takto

$$\begin{aligned} \frac{1}{2}x_0(S_1S_3 + S_2S_3) + \frac{1}{2}y_0(S_1S_2 + S_2S_3) + \frac{1}{2}z_0(S_1S_2 + S_1S_3) \\ = x_1S_1S_2 + y_1S_1S_3 + z_1S_2S_3 \end{aligned}$$

kde $x_1 = \frac{1}{2}(y_0 + z_0) = \frac{1}{2}(1 - x_0)$ a y_1 a z_1 podobně.

Zde můžeme použít substituci $X = x + k$, kde k je konstanta. Pak

$$X_1 + k = \frac{1}{2}(1 - X_0 - k)$$

Pokud zvolíme $k = \frac{-1}{3}$, dostaneme $X_1 = \frac{-1}{2}X_0$ a dále $X_n = \left(\frac{-1}{2}\right)^n X_0$.

Z toho dostaneme $x_n = X_n - k = \left(\frac{-1}{2}\right)^n (x_0 + k) - k = \frac{1}{3} \left[1 - \left(\frac{-1}{2}\right)^n \right] + \left(\frac{-1}{2}\right)^n x_0$, což je námi požadovaný výsledek pro četnost S_1S_2 v n -té populaci.

Stejný postup aplikujeme i na četnosti S_1S_3 a S_2S_3 , tedy na y a z .

V tomto případě, pokud n bude dostatečně velké, konvergují jednotlivé genotypové četnosti k $\frac{1}{3}$.

Příklad: Uvažujme náhodné křížení s počáteční populací:

$$\frac{1}{2}S_1S_2 + \frac{1}{3}S_1S_3 + \frac{1}{6}S_2S_3$$

kde S_1 , S_2 a S_3 jsou autosterilizační geny. Následující tabulka (*Tabulka 2*) pak uvádí četnosti jednotlivých genotypů v následných generacích.

Tím je ověřena konvergence k $\frac{1}{3}$ a je zde snadno vidět, že k této rovnováze se populace výrazně blíží již v desáté generaci náhodného křížení.

Tabulka 2

Generace	Genotyp		
	S_1S_2	S_1S_3	S_2S_3
0	0,5000	0,3333	0,1667
5	0,3073	0,3125	0,3177
10	0,3341	0,3340	0,3338
50	0,3333	0,3333	0,3333
100	0,3333	0,3333	0,3333
1000	0,3333	0,3333	0,3333

Obecný případ s více než třemi alelami není jednoduše řešitelný, ale pravděpodobně rovnováha v populaci nastane pouze tehdy, pokud budou mít všechny genotypy stejné četnosti.

2. 6 Dva lokusy, každý se dvěma alelami

Jakmile uvažujeme dva lokusy, je třeba vzít v úvahu, že mohou být vzájemně provázené. Tedy že se geny na těchto dvou lokusech vzájemně ovlivňují. Statisticky můžeme tento problém dle Kempthorna chápat takto:

1. Jedinec typu AB/ab produkuje gamety s následujícími pravděpodobnostmi:

AB	Ab	aB	ab
$\frac{1}{2}(1-r)$	$r/2$	$r/2$	$\frac{1}{2}(1-r)$

2. Jedinec typu Ab/aB produkuje gamety s následujícími pravděpodobnostmi:

AB	Ab	aB	ab
$r/2$	$\frac{1}{2}(1-r)$	$\frac{1}{2}(1-r)$	$r/2$

Číslo r je nazýváno rekombinační hodnota a je to podíl rekombinovaných (cross-over)² gamet. Celkem zde máme devět genotypů se dvěma lokusy a dvěma alelami na každém lokusu, ale díky rozdílné produkci gamet mezi dvěma typy dvojitých heterozygot (jak bylo výše uvedeno), máme vlastně deset typů jedinců, což ukazují i výsledky v chovech. Označíme frekvenci genotypů v n -té generaci jako $f_{ij}^{(n)}$, kde i je rovno 0, 1 nebo 2 v závislosti na počtu přítomných A genů a j stejně tak pro B . Tedy četnost $Aabb$ v n -té generaci bude zapsána jako $f_{10}^{(n)}$. Pro četnosti párujících (AB/ab) a odmítajících se (Ab/aB) heterozygot použijeme značení $f_{11c}^{(n)}$ a $f_{11r}^{(n)}$.

Vývoj populace dle náhodného křížení je založen na naší základní větě – genotypové pole generace n je čtvercem gametických polí generace $n - 1$. Následující tabulka (Tabulka 3) představuje deset genotypů, jejich četnosti v generaci 0 a gamety vyprodukované touto populací.

2 Crossing-over (rekombinační proces) je proces při dělení buňky, ve kterém si mohou jednotlivé chromozomy vyměňovat své nesesterské úseky – tedy úseky jednoho a druhého párového chromozomu. Tak vzniknou gamety s nerekombinovanými (AB/ab) a s rekombinovanými (Ab/aB) genotypy. Rekombinované gamety však vzniknou pouze tehdy, pokud mezi lokusem nesoucím pár A/a a pár B/b dojde ke crossing-overu [2]. Rekombinační hodnota tudíž vyjadřuje četnost výskytu crossing-overu.

Tabulka 3

Genotyp	Četnost v generaci 0	Četnosti vyprodukovaných gamet			
		<i>AB</i>	<i>Ab</i>	<i>aB</i>	<i>ab</i>
<i>AABB</i>	$f_{22}^{(0)}$	1	0	0	0
<i>AABb</i>	$f_{21}^{(0)}$	½	½	0	0
<i>AAbb</i>	$f_{20}^{(0)}$	0	1	0	0
<i>AaBB</i>	$f_{12}^{(0)}$	½	0	½	0
<i>AB/ab</i>	$f_{11c}^{(0)}$	½ (1- <i>r</i>)	½ <i>r</i>	½ <i>r</i>	½ (1- <i>r</i>)
<i>Ab/aB</i>	$f_{11r}^{(0)}$	½ <i>r</i>	½ (1- <i>r</i>)	½ (1- <i>r</i>)	½ <i>r</i>
<i>Aabb</i>	$f_{10}^{(0)}$	0	½	0	½
<i>aaBB</i>	$f_{02}^{(0)}$	0	0	1	0
<i>aaBb</i>	$f_{01}^{(0)}$	0	0	½	½
<i>aabb</i>	$f_{00}^{(0)}$	0	0	0	1

Gametické pole populace je tedy

$$\begin{aligned}
& (f_{22}^{(0)} + \frac{1}{2}f_{21}^{(0)} + \frac{1}{2}f_{12}^{(0)} + \frac{1}{2}(1-r)f_{11c}^{(0)} + \frac{1}{2}rf_{11r}^{(0)})AB \\
& + (\frac{1}{2}f_{21}^{(0)} + f_{20}^{(0)} + \frac{1}{2}rf_{11c}^{(0)} + \frac{1}{2}(1-r)f_{11r}^{(0)} + \frac{1}{2}f_{10}^{(0)})Ab \\
& + (\frac{1}{2}f_{12}^{(0)} + \frac{1}{2}rf_{11c}^{(0)} + \frac{1}{2}(1-r)f_{11r}^{(0)} + f_{02}^{(0)} + \frac{1}{2}f_{01}^{(0)})aB \\
& + (\frac{1}{2}(1-r)f_{11c}^{(0)} + \frac{1}{2}rf_{11r}^{(0)} + \frac{1}{2}f_{10}^{(0)} + \frac{1}{2}f_{01}^{(0)} + f_{00}^{(0)})ab,
\end{aligned}$$

což můžeme označit jako $P_{11}^{(0)}AB + P_{10}^{(0)}Ab + P_{01}^{(0)}aB + P_{00}^{(0)}ab$.

Jak jsme již viděli v kapitole 2. 2, populace bude v rovnováze pro dva oddělené lokusy po jedné generaci náhodného křížení.

Z daného gametického pole můžeme odvodit genotypové a dále i gametické pole generace 1. Například řekněme, že četnost jedinců *AABB* je $(P_{11}^{(0)})^2$, *AB/ab* je $2P_{11}^{(0)}P_{00}^{(0)}$ atd. Četnost *AB* gamet pak bude

$$P_{11}^{(0)2} + P_{11}^{(0)}P_{10}^{(0)} + P_{11}^{(0)}P_{01}^{(0)} + P_{11}^{(0)}P_{00}^{(0)}(1-r) + rP_{10}^{(0)}P_{01}^{(0)},$$

což je za podmínky $P_{11}^{(0)} + P_{10}^{(0)} + P_{01}^{(0)} + P_{00}^{(0)} = 1$

rovno $P_{11}^{(0)} - r(P_{11}^{(0)}P_{00}^{(0)} - P_{10}^{(0)}P_{01}^{(0)}) = P_{11}^{(0)} - r\Delta^{(0)}$,

kde $\Delta^{(0)} = P_{11}^{(0)}P_{00}^{(0)} - P_{10}^{(0)}P_{01}^{(0)}$.

Pokud tento postup uplatníme i pro ostatní typy gamet, dostaneme gametické pole generace 1 jako

$$P_{11}^{(1)}AB + P_{10}^{(1)}Ab + P_{01}^{(1)}aB + P_{00}^{(1)}ab,$$

kde $P_{11}^{(1)} = P_{11}^{(0)} - r\Delta^{(0)}$, $P_{10}^{(1)} = P_{10}^{(0)} + r\Delta^{(0)}$,

$$P_{01}^{(1)} = P_{01}^{(0)} + r\Delta^{(0)}, \quad P_{00}^{(1)} = P_{00}^{(0)} - r\Delta^{(0)}.$$

Necht' $\Delta^{(1)} = P_{11}^{(1)}P_{00}^{(1)} - P_{10}^{(1)}P_{01}^{(1)}$,

$$\begin{aligned}
\Delta^{(1)} &= (P_{11}^{(0)} - r\Delta^{(0)})(P_{00}^{(0)} - r\Delta^{(0)}) - (P_{10}^{(0)} + r\Delta^{(0)})(P_{01}^{(0)} + r\Delta^{(0)}) \\
&= (P_{11}^{(0)}P_{00}^{(0)} - P_{10}^{(0)}P_{01}^{(0)}) - r\Delta^{(0)}(P_{11}^{(0)} + P_{10}^{(0)} + P_{01}^{(0)} + P_{00}^{(0)}) + r^2\Delta^{(0)2} - r^2\Delta^{(0)2},
\end{aligned}$$

což je za podmínky $P_{11}^{(0)} + P_{10}^{(0)} + P_{01}^{(0)} + P_{00}^{(0)} = 1$ rovno $\Delta^{(0)} - r \Delta^{(0)}$.

Podobně $\Delta^{(2)} = \Delta^{(1)} - r \Delta^{(1)}$, tedy $\Delta^{(n)} = (1 - r)^n \Delta^{(0)}$.

A z toho pokud n jde k nekonečnu a r je mezi 0 a 1, $\Delta^{(n)}$ jde k nule.

Nyní najděme $P_{11}^{(n)}$. Máme

$$P_{11}^{(n)} - P_{11}^{(n-1)} = -r \Delta^{(n-1)},$$

.....

$$P_{11}^{(1)} - P_{11}^{(0)} = -r \Delta^{(0)}.$$

Dosazením dostaneme

$$\begin{aligned} P_{11}^{(n)} - P_{11}^{(0)} &= -r (\Delta^{(0)} + \Delta^{(1)} + \dots + \Delta^{(n-1)}) \\ &= -r \Delta^{(0)} (1 + (1-r) + (1-r)^2 + \dots + (1-r)^{n-1}) \\ &= -\Delta^{(0)} (1 - (1-r)^n). \end{aligned}$$

Tudíž máme výsledek $P_{11}^{(n)}$ konverguje k $P_{11}^{(0)} - \Delta^{(0)} = P_{11}$,
 podobně $P_{10}^{(n)}$ konverguje k $P_{10}^{(0)} - \Delta^{(0)} = P_{10}$,
 $P_{01}^{(n)}$ konverguje k $P_{01}^{(0)} - \Delta^{(0)} = P_{01}$,
 $P_{00}^{(n)}$ konverguje k $P_{00}^{(0)} - \Delta^{(0)} = P_{00}$.

Rozdíl mezi četnostmi AB/ab a Ab/aB heterozygot by měl být

$$2(P_{11}^{(n)}P_{00}^{(n)} - P_{10}^{(n)}P_{01}^{(n)}) = 2 \Delta^{(n)}.$$

Jelikož jsme však zjistili, že v případě velkého n směřuje $\Delta^{(n)}$ k nule, četnosti těchto heterozygot jsou si rovné.

Genotypové pole je tedy dáno jako

$$(P_{11}AB + P_{10}Ab + P_{01}aB + P_{00}ab)^2.$$

Jelikož $P_{11}P_{00} - P_{10}P_{01} = 0$,

$P_{11} = f_A f_B$, $P_{10} = f_A f_b$, $P_{01} = f_a f_B$, $P_{00} = f_a f_b$, kde f jsou četnosti genů.

Máme tedy konečný výsledek, že chování náhodně se křížících populací se dvěma lokusy se blíží populacím, jejichž genotypové četnosti jsou dány

$$((f_A A + f_a a)(f_B B + f_b b))^2,$$

kde f_A, f_a, f_B, f_b jsou původní četnosti genů.

2.7 Obecné chování dvou provázaných lokusů

Nyní uvedeme obecné pravidlo pro případ, kdy je počet alel libovolný pro každý lokus, ale nemusí být vždy stejný pro oba lokusy.

Ze základní věty vyplývá, že to jediné, čím může prvotní generace ovlivnit svou následující, je její gametické pole. Písmenem A s indexem i (i_1 či i_2) označme alely v jednom lokusu a písmenem B s indexem j (j_1, j_2) alely z druhého lokusu. Genotyp bude označen jako $A_{i_1} A_{i_2} B_{j_1} B_{j_2}$, kde se předpokládá, že jedinec s tímto genotypem získal geny A_{i_1} a B_{j_1} od otce a geny $A_{i_2} B_{j_2}$ od matky. Vazba bude pak produkovat následující rozdělení gamet

$$\begin{array}{cccc} A_{i_1} B_{j_1} & A_{i_2} B_{j_1} & A_{i_1} B_{j_2} & A_{i_2} B_{j_2} \\ \frac{1}{2}(1-r) & \frac{1}{2}r & \frac{1}{2}r & \frac{1}{2}(1-r) \end{array}$$

Poznamenejme ještě, že pokud $i_1 = i_2 = i$, budou vytvářeny pouze dva typy gamet – $A_i B_{j_1}$ a $A_i B_{j_2}$ budou shodné stejně jako $A_{i_1} B_{j_2}$ a $A_{i_2} B_{j_1}$ – a tyto budou vytvářeny ve stejných počtech. Necht' gametické pole generace 0 je $\sum_{ij} P_{ij} A_i B_j$.

Následující generace bude mít genotypové pole

$$\left(\sum_{i_1 j_1} P_{i_1 j_1} A_{i_1} B_{j_1}\right) \left(\sum_{i_2 j_2} P_{i_2 j_2} A_{i_2} B_{j_2}\right) = \sum_{i_1 j_1 i_2 j_2} P_{i_1 j_1} P_{i_2 j_2} A_{i_1} B_{j_1} A_{i_2} B_{j_2},$$

což je v souladu s tím, že jedinec v této generaci získá náhodnou gametu jak od otce, tak od matky. Když použijeme výše uvedené četnosti gamet, je vidět, že gametické pole generace 1 je

$$\begin{aligned} & \sum_{i_1 i_2 j_1 j_2} P_{i_1 j_1} P_{i_2 j_2} \left[\frac{1}{2}(1-r) A_{i_1} B_{j_1} + \frac{1}{2}r A_{i_2} B_{j_1} + \frac{1}{2}r A_{i_1} B_{j_2} + \frac{1}{2}(1-r) A_{i_2} B_{j_2} \right] \\ &= \frac{1}{2}(1-r) \sum_{i_1 i_2 j_1 j_2} P_{i_1 j_1} P_{i_2 j_2} A_{i_1} B_{j_1} + \frac{1}{2}r \sum_{i_1 i_2 j_1 j_2} P_{i_1 j_1} P_{i_2 j_2} A_{i_2} B_{j_1} \\ & \quad + \frac{1}{2}r \sum_{i_1 i_2 j_1 j_2} P_{i_1 j_1} P_{i_2 j_2} A_{i_1} B_{j_2} + \frac{1}{2}(1-r) \sum_{i_1 i_2 j_1 j_2} P_{i_1 j_1} P_{i_2 j_2} A_{i_2} B_{j_2} \end{aligned}$$

Po přeindexování (ve druhé sumě vzájemně vyměníme i_1 s i_2 , ve třetí j_1 s j_2 a ve čtvrté i_1 s i_2 a j_1 s j_2 , což si můžeme dovolit díky nezávislosti proměnných) dostaneme gametické pole generace 1 v tomto tvaru

$$\begin{aligned} & \sum_{i_1 j_1} \sum_{i_2 j_2} \left[\frac{1}{2}(1-r) P_{i_1 j_1} P_{i_2 j_2} + \frac{1}{2}r P_{i_2 j_1} P_{i_1 j_2} + \frac{1}{2}r P_{i_1 j_2} P_{i_2 j_1} + \frac{1}{2}(1-r) P_{i_2 j_2} P_{i_1 j_1} \right] A_{i_1} B_{j_1} \\ &= \sum_{i_1 j_1} P'_{i_1 j_1} A_{i_1} B_{j_1} \end{aligned} \quad (1)$$

Co ale $P'_{i_1 j_1}$ ve skutečnosti je?

Označme $\sum_{i_2 j_2} P_{i_2 j_2} = 1$, $\sum_{j_2} P_{i_1 j_2} = F_{i_1}$, $\sum_{i_2} P_{i_2 j_1} = G_{j_1}$, kde G_{j_1} je četnost genu B_{j_1} na lokusu B a F_{i_1} je četnost genu A_{i_1} na lokusu A .

Máme tedy

$$\begin{aligned} P'_{i_1 j_1} &= \frac{1}{2}(1-r) P_{i_1 j_1} + \frac{1}{2}r G_{j_1} F_{i_1} + \frac{1}{2}r F_{i_1} G_{j_1} + \frac{1}{2}(1-r) P_{i_1 j_1} \\ &= (1-r) P_{i_1 j_1} + r F_{i_1} G_{j_1} = P_{i_1 j_1} + r(F_{i_1} G_{j_1} - P_{i_1 j_1}) \end{aligned}$$

Nyní dokážeme fakt, že při náhodném křížení se četnost genů nemění z generace na generaci. Pro generaci 1 tedy máme $F'_{i_1} = \sum_{j_1} P'_{i_1 j_1} = F_{i_1} + r(F_{i_1} - F_{i_1}) = F_{i_1}$ a podobně $G'_{j_1} = G_{j_1}$.

Pokud z (1) na levé straně ubereme $F'_{i_1} G'_{j_1}$ a na pravé $F_{i_1} G_{j_1}$, stále máme rovnost.

$$\begin{aligned}
P'_{i_1j_1} - F'_{i_1}G'_{j_1} &= P_{j_1j_1} + r(F_{i_1}G_{j_1} - P_{i_1j_1}) - F_{i_1}G_{j_1} \\
&= (1-r)P_{i_1j_1} - (1-r)F_{i_1}G_{j_1} \\
&= (1-r)(P_{i_1j_1} - G_{i_1}G_{j_1})
\end{aligned}$$

Označme $P_{i_1j_1} - G_{i_1}G_{j_1}$ jako Δ_{ij} . Tímto získáme jednoduchou rovnost

$$\Delta'_{ij} = (1-r)\Delta_{ij}.$$

Můžeme tedy říct, že Δ_{ij} klesá v poměru s $(1-r)$ pro každou generaci. Dle značení v předchozí kapitole máme $\Delta_{ij}^{(n)} = (1-r)^n \Delta_{ij}^{(0)}$ a můžeme zapsat

$$P_{ij}^{(n)} = F_iG_j + (1-r)^n \Delta_{ij}^{(0)},$$

z čehož lze spočítat genotypové pole jakékoli generace.

Jelikož r je mezi 0 a 1, $\Delta^{(n)}$ klesá k nule při velkém n a gametické pole směřuje k

$$\sum_{ij} F_iG_jA_iB_j = \left(\sum_i F_iA_i\right)\left(\sum_j G_jB_j\right),$$

což se rovná gametickému poli populace s genotypovým polem

$$\sum_{ijmn} F_iF_jG_mG_nA_iA_jB_mB_n,$$

anebo populaci, kde jsou všechny geny kombinovány náhodně striktně dle jejich četností.

Na konec poznamenejme, že r je zásadně omezeno na interval mezi 0 a $\frac{1}{2}$ (Fisher, 1948). Pokud je r rovno $\frac{1}{2}$, získáme předešlý výsledek pro dvě neprovázané alely. Pokud je si provázanost málo blízká, rekombinant r je blízko nuly a populace dosahuje rovnováhy velmi pomalu.

3 Základní problémy výběru

Prozkoumali jsme jednoduchou ukázkou populační dynamiky – náhodné křížení. Náhodné křížení, jak jsme již zmínili, obsahuje více podmínek než jen tu zřejmou, že jedinci se páří náhodně v populaci. Nejdůležitější podmínkou pro izolované populace je ta, že úmrtnost je stejná pro všechny genotypy; každý jedinec a každé páření vyprodukuje stejný počet potomků. Naproti tomu si můžeme představit situaci v reálné populaci, kde spáření $AA \times AA$ zplodí určitý počet potomků, zatímco jiné páření $Aa \times aa$ pouze poloviční počet potomků.

Často doufáme, že volné opylení vede k náhodnému křížení. To ale nikdy nenastane, dokud všechny rostliny nevyprodukují stejné množství pylu, který pak bude za pomoci větru (či jiných sil) roznesen rovnoměrně na všechny rostliny a samozřejmě produkce pylu musí nastat ve stejném časovém období. Nebudeme se zde však zabývat kompletním pokrytím, ale pouze několika jednoduššími příklady selekce.

3.1 Výběr založený na genotypu

Uvažujme případ, kdy se generace nepřekrývají. Je zde mnoho variant, které mohou nastat, ale v podstatě všechny vedou ke stejnému matematickému modelu. Například můžeme mít umělý výběr, kde se chovatel snaží o nenáhodné párování založené na genotypech, nebo situaci, ve které se vyskytuje náhodné křížení jedinců, ale schopnost reprodukce je vázaná na genotyp.

Vezměme tedy případ kde je schopnost přežití jedinců – schopnost reprodukce – vázaná na genotyp. Tedy každý jedinec má jiné předpoklady k přežití. Nechť prvotní populace při narození je $p^2 AA + 2pqAa + q^2 aa$ a nechť schopnost přežití je vyjádřena poměrem $1 - S : 1 : 1 - s$ pro genotypy AA , Aa a aa . Pak populace dospělých (přeživších) jedinců má dle Kempthorna následující genotyp:

$$\frac{(1 - S)p^2 AA + 2pqAa + (1 - s)q^2 aa}{1 - Sp^2 - sq^2} .$$

Populace pak produkuje tyto gamety $\frac{(p - Sp^2)A + (q - sq^2)a}{1 - Sp^2 - sq^2}$ a četnost genu A

při narození v další generaci je $\frac{p - Sp^2}{1 - Sp^2 - sq^2} = p^{(1)}$, tedy

$$p^{(1)} - p = \frac{-pq(Sp - sq)}{1 - Sp^2 - sq^2} = -pq \frac{S - (S + s)q}{1 - Sp^2 - sq^2} .$$

Rovnováha tedy existuje a je stabilní pokud $q = \frac{S}{S + s}$ a zároveň S i s se nacházejí v intervalu od 0 do 1.

S tímto obecným vztahem můžeme přistoupit ke konkrétnějšímu příkladu. Příklad, který je v zájmu chovatelů je ten, kdy $S = 0$ a $s = 1$. To nastává tehdy, když dominantní homozygot a heterozygot mají stejnou úmrtnost a stejnou schopnost reprodukce, zatímco recesivní jsou eliminováni či sterilizováni. V tomto případě máme

$$p^{(1)} = \frac{p}{1 - q^2} = \frac{p}{(1 - q)(1 + q)} = \frac{1}{1 + q} , \quad q^{(1)} = 1 - p^{(1)} = 1 - \frac{1}{1 + q} = \frac{q}{1 + q} ,$$

kde $p^{(1)}$ je četnost genu A a $q^{(1)}$ genu a v první generaci.

Je snadné popsat vývoj populace tímto systémem výběru, pro který máme $\frac{1}{q^{(1)}} = \frac{1+q}{q} = \frac{1}{q} + 1$, což vede k $\frac{1}{q^{(n)}} = \frac{1}{q} + n$, kde $q^{(n)}$ je četnost a v n -té generaci. Přímý vztah pro $q^{(n)}$ je

$$q^{(n)} = \frac{q}{1 + nq}.$$

$q^{(n)}$ tedy spěje k nule, ale velmi pomalu, pokud je q malé.

Příklad: Uvažujme teoretický příklad albinismu v králičí populaci, který se projevuje jen u recesivních jedinců, zatímco u jedinců dominantních a dominantně heterozygotních se tento znak neprojeví. Chovatel chce mít svůj chov prostý albínů, a proto tyto jedince nepřipouští k páření. Konkrétně nás nyní bude zajímat, za kolik náhodně křížících se generací s vyloučenými albíny se změní jejich četnost v populaci z 1 : 1 000 na 1 : 1 000 000.

Zde stačí pouze vyjádřit $n = \frac{q - q^{(n)}}{q^{(n)}q}$, kde $q = 1000$ a $q^{(n)} = 1\,000\,000$.

Odtud dostaneme, že tisíckrát se četnost albínů v králičí populaci zmenší až za 999 000 generací.

Další ukázkou je situace, kde jsou veškerá heterozygota redukována ve prospěch homozygotních jedinců – tedy pouze homozygotní jedinci jsou schopni reprodukce.

Předpokládejme, že prvotní populace je následující:

$$\begin{array}{ll} AA & aa \\ P & R \quad P + R = 1 \end{array}$$

Z toho můžeme získat genotypové pole jedinců při narození v generaci 1 $P^2 AA + 2PRAa + R^2 aa$ a dospělých $\frac{P^2 AA + R^2 aa}{P^2 + R^2}$, kde $P^2 + R^2 = 1 - 2PR$.

Z tohoto dostaneme genotypové pole dospělých v generaci 2:

$$\frac{P^4 AA + R^4 aa}{(P^2 + R^2)^2 [1 - 2P^2 R^2 / (P^2 + R^2)^2]} \text{ a tak dále.}$$

Četnost genu a při narození je tedy následující:

Generace	Četnost
0	R
1	$\frac{R^2}{P^2 + R^2}$
2	$\frac{R^4}{(P^2 + R^2)^2} / [1 - \frac{2P^2 R^2}{(P^2 + R^2)^2}]$

Takže pokud $q^{(i)}$ představuje četnost genu a v generaci i , máme

$$q^{(i+1)} = \frac{[q^{(i)}]^2}{1 - 2q^{(i)} + 2[q^{(i)}]^2} .$$

To dává $\frac{1}{q^{(i+1)}} = \frac{1}{[q^{(i)}]^2} - \frac{2}{q^{(i)}} + 2$ neboli $\frac{1}{q^{(i+1)}} - 1 = \left[\frac{1}{q^{(i)}} - 1\right]^2$.

Odtud máme obecně $\frac{1}{q^{(n)}} - 1 = \left[\frac{1}{q^{(0)}} - 1\right]^{2^n}$ neboli $q^{(n)} = \frac{[q^{(0)}]^{2^n}}{[p^{(0)}]^{2^n} + [q^{(0)}]^{2^n}}$.

Limitní chování, když n jde k nekonečnu, můžeme uvést vztahem

$$\log\left(\frac{1 - q^{(n)}}{q^{(n)}}\right) = 2^n \log\left(\frac{1 - q^{(0)}}{q^{(0)}}\right) .$$

Pokud je $q^{(0)}$ menší než $\frac{1}{2}$, pak $\frac{1 - q^{(0)}}{q^{(0)}}$ je větší než 1, takže $\log\left(\frac{1 - q^{(n)}}{q^{(n)}}\right) \rightarrow \infty$, což může nastat tehdy a jen tehdy, pokud $q^{(n)}$ jde k nule. Pokud je $q^{(0)}$ větší než $\frac{1}{2}$, pak $\frac{1 - q^{(0)}}{q^{(0)}}$ je menší než 1, $\log\left(\frac{1 - q^{(0)}}{q^{(0)}} je záporný a $\log\left(\frac{1 - q^{(n)}}{q^{(n)}}\right) \rightarrow -\infty$ nebo $q^{(n)}$ jde k 1.$

Tento systém selekce nám ukazuje, že čím rozšířenější gen je prvotně zahrnut, tím je populace zdravější – neporušenější. Pokud bude $q^{(0)}$ přesně $\frac{1}{2}$, q se nebude měnit z generace na generaci, ale zmíněné argumenty ukazují, že jakákoli odchylka z této rovnováhy bude mít za následek pohyb generace ve směru tohoto odchýlení. Rovnováha při $q^{(0)} = \frac{1}{2}$ je tedy velmi nestabilní.

3.2 Výběr omezený na pohlaví

Zde uvažujme například selekci autozomálních³ genů, kde budou vybírání pouze samci. Jako příklad uvedeme situaci, kdy mohou být recesivní samci zcela odstraněni. Nechť gametické pole produkované samci v generaci n je $u_n A + (1 - u_n)a$ a samicemi $v_n A + (1 - v_n)a$.

Pak se generace $n + 1$ při narození skládá z

$$u_n v_n A A + (u_n(1 - v_n) + v_n(1 - u_n))A a + (1 - u_n)(1 - v_n)a a .$$

Reprodukční část samčí populace se po odvržení všech recesivních jedinců skládá z $\{u_n v_n A A + [u_n(1 - v_n) + v_n(1 - u_n)]A a\} / [1 - (1 - u_n)(1 - v_n)]$ s gametickým polem

$$\left[\left(\frac{u_n + v_n}{2}\right)A + \left(\frac{u_n(1 - v_n) + v_n(1 - u_n)}{2}\right)a\right] / [1 - (1 - u_n)(1 - v_n)] .$$

Samičí gametické pole je $\left(\frac{u_n + v_n}{2}\right)A + \left(\frac{2 - u_n - v_n}{2}\right)a$.

3 Autozom (somatický) chromozom je každý jiný chromozom než pohlavní (X a Y). Tedy autozomální (somatické) geny jsou geny nevztahující se k pohlaví jedince [2].

Z toho tedy máme $v_{n+1} = \frac{u_n + v_n}{2}$, $u_{n+1} = \frac{u_n + v_n}{2} \frac{1}{1 - (1 - u_n)(1 - v_n)}$,
 $v_{n+2} = \frac{u_{n+1} + v_{n+1}}{2} = \frac{u_n + v_n}{4} \left[1 + \frac{1}{1 - (1 - u_n)(1 - v_n)} \right]$, takže

$v_{n+2} = \frac{1}{2} v_{n+1} \left[1 + \frac{1}{1 - (1 - v_n)(1 + v_n + 2v_{n+1})} \right]$, což je rekurentní vyjádření četnosti genu A v samičí populaci. Toto rekurentní vyjádření je nelineární a jednoduché řešení nejspíš neexistuje.

3.3 Selektce pohlavně vázaných genů

Zde z velkého množství možností vyberme příklad, ve kterém uvažujeme situaci, kdy jsou odvrženi všichni recesivní jedinci. Necht' je při narození prvotní populace následující:

Samice				Samci
AA	Aa	aa	A	a
$r^{(0)}$	$2s^{(0)}$	$t^{(0)}$	$u^{(0)}$	$v^{(0)}$

Pak po úplném odmítnutí všech recesivních jedinců je samičí produkce gamet $((r^{(0)} + s^{(0)})A + s^{(0)}a) / (1 - t^{(0)})$ a u samců je to pouze A .

Tudíž v následující generaci budou jednotlivé četnosti zastoupení genotypů následující:

$$u^{(1)} = \frac{r^{(0)} + s^{(0)}}{1 - t^{(0)}}, \quad v^{(1)} = \frac{s^{(0)}}{1 - t^{(0)}} ,$$

$$r^{(1)} = \frac{r^{(0)} + s^{(0)}}{1 - t^{(0)}}, \quad 2s^{(1)} = \frac{s^{(0)}}{1 - t^{(0)}}, \quad t^{(1)} = 0 .$$

Budeme-li pokračovat stejným způsobem dále, dostaneme četnosti v generaci 2

$$u^{(2)} = r^{(1)} + s^{(1)}, \quad v^{(2)} = s^{(1)},$$

$$r^{(2)} = r^{(1)} + s^{(1)}, \quad 2s^{(2)} = s^{(1)}, \quad t^{(2)} = 0 .$$

Tudíž recesivní samice jsou zcela eliminovány v jedné generaci a po jedné generaci jsou samičí heterozygoty redukovány na jednu polovinu v každé další generaci.

Pokud je selektce pohlavně vázaná, tudíž jsou například odvrhováni pouze recesivní samci, v generaci 0 se gametická produkce samic skládá z $(r^{(0)} + s^{(0)})A + (s^{(0)} + t^{(0)})a$ a samců pouze z A .

Tedy samičí populace v generaci 1 má genotypové pole $(r^{(0)} + s^{(0)})AA + (s^{(0)} + t^{(0)})Aa$ a samčí $(r^{(0)} + s^{(0)})A + (s^{(0)} + t^{(0)})a$.

Necht' $p^{(i)} = r^{(i)} + s^{(i)}$ a $q^{(i)} = 1 - p^{(i)}$, dostaneme

$$p^{(1)} = r^{(0)} + s^{(0)} + \frac{1}{2} (s^{(0)} + t^{(0)})$$

$$= p^{(0)} + \frac{1}{2} q^{(0)}$$

$$= p^{(0)} + \frac{1}{2} (1 - p^{(0)})$$

$$= \frac{1}{2} + \frac{1}{2} p^{(0)},$$

což je četnost AA samic v generaci 1 a $q^{(1)} = \frac{1}{2} q^{(0)}$, tedy $q^{(n)} = (\frac{1}{2})^n q^{(0)}$, tj. četnost Aa

samic v generaci n . Četnost heterozygotních samic tudíž klesá o jednu polovinu v každé generaci.

Příklad: Uvažujme pohlavně vázaný znak, k jehož selekci dochází pouze u homogametického pohlaví – tedy samičího – a dále předpokládejme, že podíl k recesivních samic je odvržen.

Máme tedy následující populaci:

	Genotypové pole	Gametické pole
Samice	$rAA + 2sAa + taa$	$(r + s)A + (s + (1 - k)t)a / (1 - kt)$
Samci	$uA + va$	$uA + va$

Genotypové pole samic v následující generaci tudíž je

$$\frac{u(r + s)AA + (v(r + s) + u(s + (1 - k)t))Aa + v(s + (1 - kt))aa}{1 - kt}$$

a samců $\frac{(r + s)A + (s + (1 - k)t)a}{1 - kt}$.

Z toho máme $u_1 = \frac{r + s}{1 - kt}$, $v_1 = \frac{s + (1 - k)t}{1 - kt}$,

$$r_1 = \frac{u(r + s)}{1 - kt}, \quad 2s_1 = \frac{v(r + s) + u(s + (1 - k)t)}{1 - kt}, \quad t_1 = \frac{v(s + (1 - k)t)}{1 - kt},$$

tedy $u_n = \frac{r_{n-1} + s_{n-1}}{1 - kt_{n-1}}$, $v_n = \frac{s_{n-1} + (1 - k)t_{n-1}}{1 - kt_{n-1}}$,

$$r_n = u_{n-1}u_n, \quad 2s_n = v_{n-1}u_n + u_{n-1}v_n, \quad t_n = v_{n-1}v_n.$$

Nyní uvažujme (opět pouze teoreticky) populaci octomilek, kde jsou takto odvrhovány samice s bílými očima. Bílou barvu očí uvažujeme podmíněnou recesivní alelou a , zatímco červenou barvu očí dominantní alelou A . Následující tabulka (*Tabulka 4*) ukazuje vývoj v populaci s počátečním samičím genotypem $0,3AA + 0,2Aa + 0,5aa$ a samčím $0,4A + 0,6a$ při dvou různých hodnotách konstanty k ($k = 0,3$ a $k = 0,6$).

Tabulka 4

<i>Generace</i>	<i>r</i>	<i>2s</i>	<i>t</i>	<i>u</i>	<i>v</i>	<i>k</i>
1	0,3	0,2	0,5	0,4	0,6	0,3
2	0,188	0,494	0,318	0,471	0,529	0,3
3	0,226	0,499	0,275	0,481	0,519	0,3
4	0,250	0,501	0,250	0,519	0,481	0,3
5	0,280	0,498	0,221	0,540	0,460	0,3
6	0,306	0,495	0,199	0,567	0,433	
<hr/>						
1	0,3	0,2	0,5	0,4	0,6	0,6
2	0,229	0,514	0,257	0,571	0,429	0,6
3	0,328	0,489	0,182	0,574	0,426	0,6
4	0,369	0,479	0,152	0,643	0,357	0,6
5	0,431	0,451	0,118	0,670	0,330	0,6
6	0,473	0,430	0,097	0,706	0,294	

Je-li $k = 0,3$, převažují v populaci samic dominantně heterozygotní jedinci a v samčí populaci ti dominantní. V šesté generaci je již 68,4 % octomilek s červenými očima. Pokud bude $k = 0,6$, mezi samicemi budou nejvíce zastoupeny dominantní stejně tak jako mezi samci – celkem tedy v šesté generaci 80,5 % červenookých octomilek.

Je zde také snadno vidět, že při zvětšujícím se k klesá také četnost zastoupení samečů s bílými očima i přes to, že selekce probíhá pouze v samičí populaci.

3. 4 Jiné případy

Existuje skutečně neomezené množství případů, které bychom zde mohli uvést. Případy již zmíněné byly postaveny na výběru založeném na genotypu každého jedince. Haldane (1923-1927) předložil odlišný typ výběru, který nazval familiární (rodinnou) selekci. Tento výběr je mezi členy stejné rodiny. Představme si jednoduchý případ, kdy je zplozeno 20 embryí, ale jen pro 10 z nich je dostatek potravy. A zde jsou pak rozdíly mezi jednotlivými embryi v jejich genotypu, který se odráží v schopnosti jejich přežití v konkurenčním prostředí.

Jiný typ výběru, zajímavý v lidské populaci beroucí v potaz *Rh* faktor, vychází v schopnost přežití jedince závislou na vztahu jeho a matčina genotypu. Konkrétní příklad *Rh* faktoru byl předložen Haldanem (1941).

***Rh* faktor** je ihned po systému AB0 nejznámějším systémem antigenů⁴. Rozlišujeme zde jedince *Rh* negativní (*Rh*-) a *Rh* pozitivní (*Rh*+). *Rh*+ se dědí dominantně, zatímco *Rh*- jedinci jsou homozygotně recesivní. V České republice je přibližně 80 % obyvatelstva *Rh* pozitivní [4]. U jedinců *Rh*- se primárně nenacházejí protilátky proti *Rh*+, ale objevují se až po imunizaci jedince *Rh* pozitivní krví – například při nevhodné transfuzi či inkompatibilním těhotenství. Při inkompatibilním těhotenství, kdy dítě zdědilo po otci *Rh*+ a matka je *Rh*-, dochází k fetální erythroblastóze. Toto je stav, kdy se matčin organismus brání plodu – jelikož matčin imunitní systém tento antigen nezná, vytváří si protilátky, které vedou k různým formám novorozenecké žloutenky, v horších případech také k nervovým poruchám. Díky placentální bariéře, kdy nedochází k mísení matčiny krve s krví plodu, bývá první těhotenství bezrizikové. Ke smísení krve dochází až při porodu – pak dochází k imunizaci matky, čímž se každé další inkompatibilní těhotenství stává mnohem rizikovějším, jelikož matka už se má „čím bránit“ proti plodu. V současné době se těmto rizikům zamezuje podáváním protilátek do 72 hodin po porodu.

Představme si populaci, ve které mají přeživší pole $uRR + 2vRr + wrr$ a dále předpokládejme náhodné křížení. Zde neuvažujeme poporodní „očkování“ matek proti imunizaci. Četnost *Rr* potomků *rr* matek bude $uw + vw = (u + v)w$.

Nechť podíl k potomků nepřežije. Pak následující generace bude mít genotypové pole $(u + v)^2RR + 2(u + v)(v + w)Rr + (v + w)^2rr - (u + v)wkRr$ neboli

$$\frac{(u + v)^2 RR + (u + v)(2v + 2w - kw) Rr + (v + w)^2 rr}{1 - k(u + v)w}$$

4 Antigeny jsou různé bílkovinné podstaty nacházející se na membránách erytrocytů (červených krvinek). Pomocí antigenů rozlišujeme různé systémy krevních skupin. Nejznámějším systémem je AB0, kde rozlišujeme čtyři základní krevní skupiny (A, B, AB, 0). Dalšími systémy jsou například *Rh*, MNSs, Lewis či P. [4]

Nechť je p četnost genu r : např. $p = v + w$. Pak četnost p' v úspěšné generaci je dána

$$p' = \frac{p - \frac{1}{2}kw(1-p)}{1 - kw(1-p)}$$

Tedy
$$\Delta p = p' - p = \frac{kw(1-p)}{1 - kw(1-p)} \left(p - \frac{1}{2} \right) .$$

Z toho vyplývá, že pokud je p větší než $\frac{1}{2}$, Δp poroste, zatímco pokud bude p menší než $\frac{1}{2}$, Δp klesne. Obecný výsledek je tedy podobný jako v případě selekce heterozygot – rovnováha je tedy nestabilní.

4 Závěr

Ukázali jsme zde základní problematiku vývoje populací prezentovanou pomocí jednoduchého modelu náhodného křížení. Náhodné křížení zahrnuje ve skutečnosti často nereálné podmínky jako například stejné počty potomků, anebo že ke křížení dochází v populaci zcela náhodně. Představením výběru jsme tedy upustili právě od toho zcela náhodného křížení – zahrnuli jsme zde preferenci mezi jedinci, tudíž že někteří jedinci jsou postupně eliminováni. Co se týče předpokladu izolace populace, ten je splnitelný, pokud například uvažujeme populaci jednoho území, kdy je pravděpodobnost smísení s jinou populací velmi malá. Ale především předpoklad vždy stejného počtu potomků je velmi náročný – řekněme téměř nesplnitelný.

Pro důkaz toho, že dědičnost je skutečně jeden mechanismus fungující pro všechny jedince stejně, jsme se museli držet všech těchto teoretických předpokladů, abychom zaručili stejné příležitosti každému jedinci. A dále jsme ukázali, že díky náhodnosti křížení je v populaci zastoupení jedinců s určitými znaky konstantní – vždy v závislosti na prvotní generaci.

Použitá literatura a software

Pro svou bakalářskou práci jsem použila výhradně software *OpenOffice.org 2.0* a následující zdroje:

- [1] Kempthorne O.: *An Introduction to Genetic Statistics*, Iowa State University Press, Ames, 1969.
- [2] Nečásek J.: *Genetika*, Scientia, s.r.o., Praha, 1993.
- [3] Jelínek J., Zicháček V.: *Biologie pro gymnázia*, Nakladatelství Olomouc, Olomouc, 2002.
- [4] Šípek A.: *Krevní skupiny a jejich dědičnost*, <http://genetika.wz.cz/skupiny.htm>
- [5] Časopis ZAHRADA – Rok na zahradě 2005: *Jak vysazovat ovocné stromy*, <http://www.raselina.cz/index.php?lok=aktualita&ID=62>