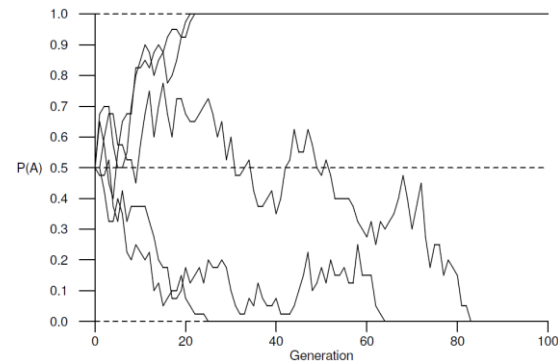
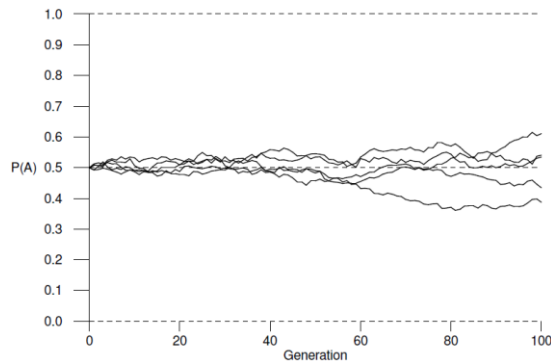


# **DRIFT II**

# EFEKTIVNÍ VELIKOST POPULACE

V reálném světě velikost populace omezená

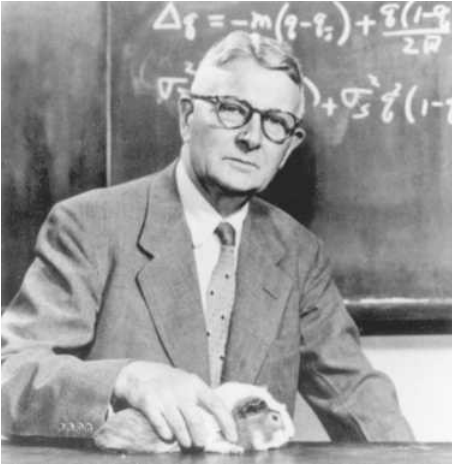
- ⇒ vyšší autozygotnost (*identity by descent*, IBD)
- ⇒ zvyšující se rozptyl frekvencí alel mezi démy v čase
- ⇒ fixace/extinkce alel



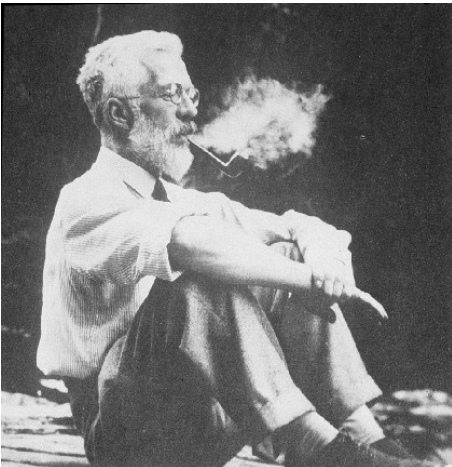
všechny tyto jevy nepřímo úměrné velikosti populace  $N$

Použitím idealizované populace můžeme odvodit přesný kvantitativní vztah driftu a velikosti populace → **Wrightův-Fisherův model**

## Wrightův-Fisherův model ještě jednou:



Sewall Wright



Ronald A. Fisher

diploidní, hermafrodit

velikost omezená, žádné fluktuační  $N$

náhodné oplození

kompletní izolace (žádný tok genů)

diskrétní generace

žádná věková struktura

žádná selekce

rozptyl výběru gamet do další generace

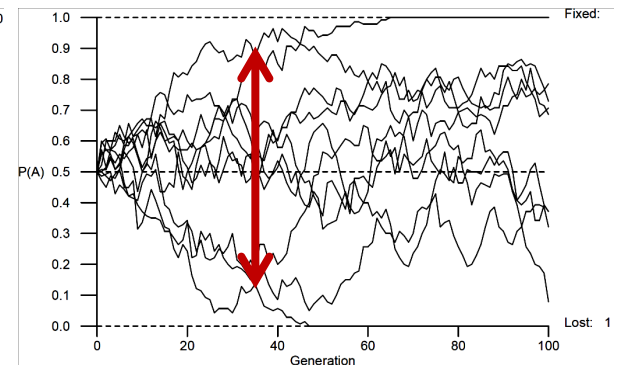
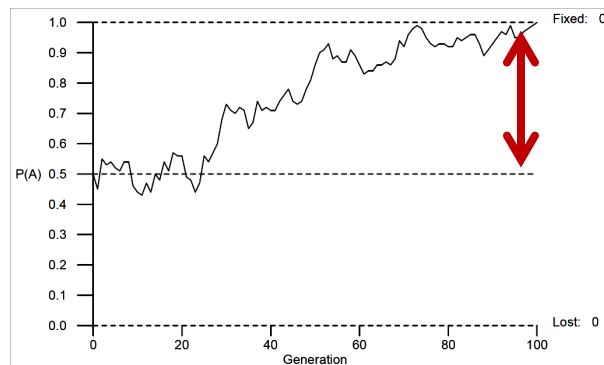
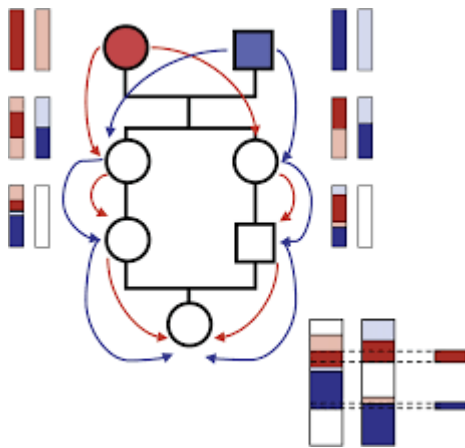
→ Poissonovo rozdělení

Efektivní velikost populace  $N_e$  nám umožňuje měřit sílu driftu v neideální populaci

## Efektivní velikost populace

= počet jedinců idealizované populace, která vykazuje stejnou míru driftu\*) jako studovaná neidealizovaná populace

\*) zde buď (1) zvýšení průměrné pravděpodobnosti autozygotnosti na autozomálním lokusu (=  $N_{eF}$ ), nebo (2) zvýšení rozptylu frekvencí alel přes generace nebo přes subpopulace (=  $N_{eV}$ )



## Podobně jako u koeficientu inbreedingu neexistuje jediná efektivní velikost!!

Díky  $N_e$  můžeme měřit sílu driftu v reálných populacích, které mohou porušovat různé předpoklady WF modelu v různé míře

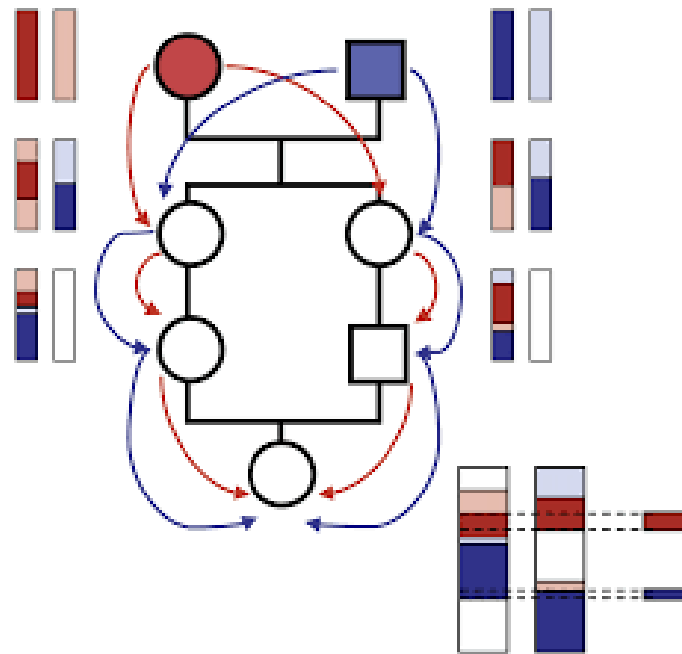
Např. efekt zakladatele:

- 1) 20 gonochoristů (ne-hermafroditů) – 10 samic, 10 samců  
jinak ~ WF model
- 2) 25 jednodomých rostlin – 50 % samosprašnost, 50 % náhodné opylení  
jinak ~ WF model

Ve kterém případě je vliv driftu silnější? →  $N_e$

# Inbreedingová efektivní velikost $N_{eF}$

vyjadřuje efektivní velikost pomocí zvýšení autozygotnosti



**Generace 0:**  $F(0) = 0$ ; velikost  $N \Rightarrow$  v každé generaci  $2N$  gamet

**Generace 1:** sice  $F(0) = 0$  (IBD by mělo být 0), ale protože jde o hermafrodity (viz WF model), můžou pocházet od stejného rodiče

Jaká je Pr., že 2 gamety pocházejí od stejného rodiče?

protože předpokládáme náhodné oplození,  $Pr = 1/N$

jestliže došlo k samooplození, Pr. že obě alely totožné =  $\frac{1}{2}$

$\Rightarrow$  celková Pr (IBD v Gen.1) =  $\frac{1}{N} \times \frac{1}{2} = \frac{1}{(2N)} = F(1)$

tj. průměrná Pr. IBD se v 1. generaci zvýšila o  $\frac{1}{(2N)}$

**Generace 2:**

$$\bar{F}(2) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) \bar{F}(1)$$

Pr. autogamie  
v Gen1

IBD v důsledku spojení gamet stejného prarodiče, vážená Pr., že tyto gamety už nejsou IBD díky autogamii

Generace  $t$ : 
$$\bar{F}(t) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) \bar{F}(t-1)$$

po několika matematických tricích: 
$$\bar{F}(t) = 1 - \left(1 - \frac{1}{2N_{eF}}\right)^t$$

a tedy

$$N_{eF} = \frac{1}{2\{1 - [1 - \bar{F}(t)]^{1/t}\}}$$

$N_{eF}$  závisí jen na prům.  $F$  a čase  $t$ , nezávisí na  $N$ !



*G. spekei*, 1979,  $N = 19$ :

$F = 0,1283$ ; prům.  $t = 1,7 \Rightarrow N_{eF} = 6,4 (\approx 1/3 N)$



kdybychom ale vzali rodiče 19 jedinců z 1979 jako referenční (nultou) generaci  $\rightarrow N_{eF}$  pro 15 potomků = 48,1 ( $\approx 2,5 N$ )!

problém v tom, že referenční generace 19 jedinců je z definice  $F = 0$

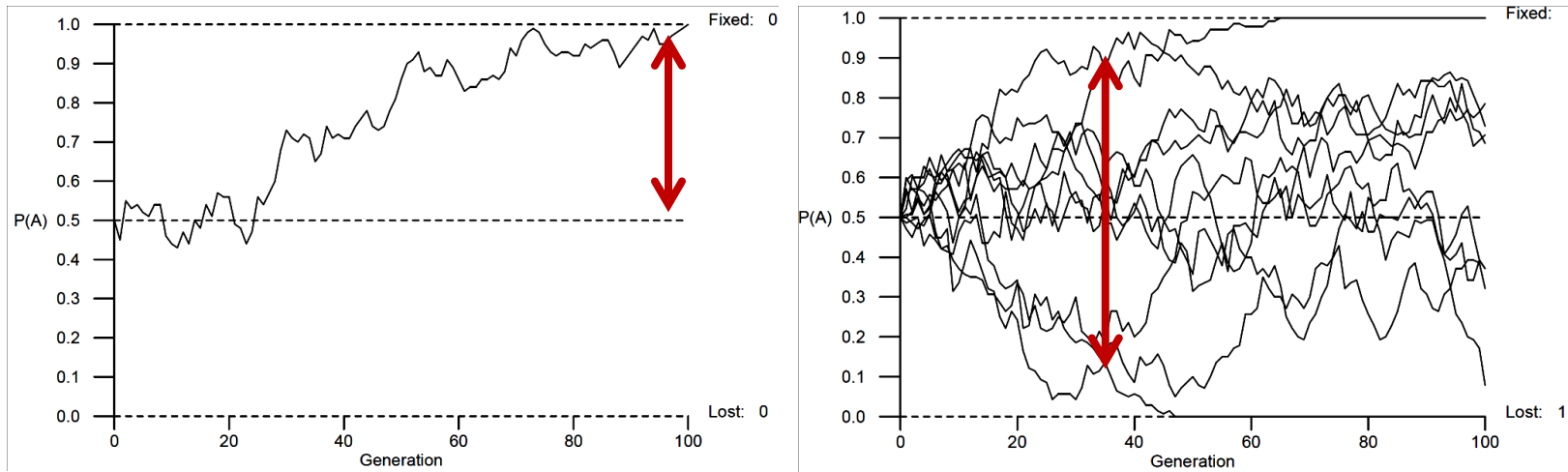
⇒ v prvním případě  $N_{eF}$  ukazuje rychlost akumulace inbreedingu po založení populace, ve druhém  $N_{eF}$  kvantifikuje účinek nového – outbredního – chovného programu



Efektivní velikost populace může být i vyšší než  $N$  (*census population size*)!

Pro stejnou populaci můžeme dostat velmi rozdílné efektivní velikosti v závislosti na volbě referenční populace!

# Varianční (rozptylová) efektivní velikost $N_{eV}$



Genetický drift způsobuje (1) náhodné odchylky od frekvencí alel v předchozí generaci a (2) rozdíly v  $p$  mezi subpopulacemi

Obojí měřeno pomocí rozptylu (variance) alelových frekvencí

Idealizovaná populace:

1 lokus: alely  $A$  o frekvenci  $p$ ,  $a$  o frekvenci  $q = p-1$

velikost populace  $N$ ,

$x$  = počet alel  $A$  ve vzorku

sampling alel v souladu s binomickým rozdělením

Protože drift nemá směr, očekávaná frekvence  $A$  v další generaci =  $p$

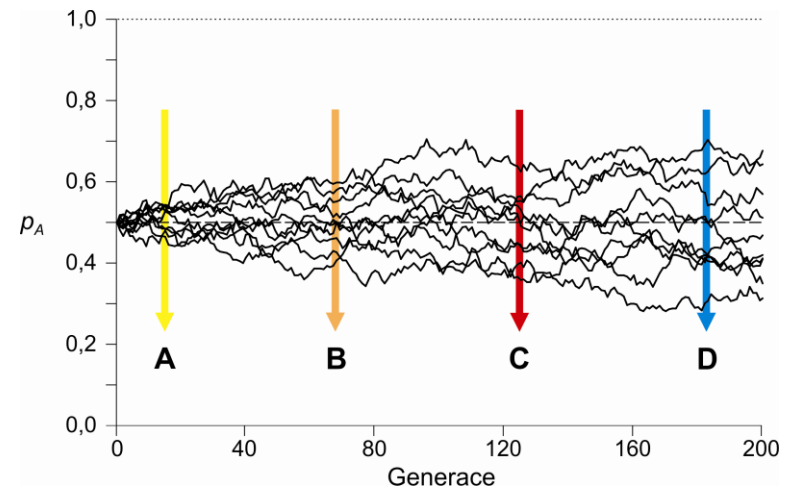
$p = x/(2N)$  = počet  $A$ /celk. počet alel

$$\text{Rozptyl alelových frekvencí} = \text{Var} \left( \frac{x}{2N} \right) = \frac{\text{Var}(x)}{(2N)^2} = \frac{2Npq}{(2N)^2} = \frac{pq}{2N}$$

v neidealizované populaci místo  $N \rightarrow N_{eV}$

Rozptyl v čase roste  $\Rightarrow$  po  $t$  generacích:

$$\text{Var}(p_t) = pq \left[ 1 - \left( 1 - \frac{1}{2N} \right)^t \right]$$



$$\text{Var}(p_t) \rightarrow \sigma_t^2$$

s tím, jak  $t \rightarrow \infty$ ,  
 $\text{Var} \rightarrow \text{Var}(\text{max}) = pq$   
 (= stav, kdy všechny alely  
 buď fixovány, nebo ztraceny)

Varianční efektivní velikost

$$N_{ev} = \frac{1}{2 \left\{ 1 - \left[ 1 - \frac{\sigma_t^2}{pq} \right]^{1/t} \right\}}$$

závisí pouze na  $\text{Var}$ ,  $p$ ,  
 $q$  a  $t$ , nezávisí na  $N$ !

frekvence  $A$  v  
 počáteční generaci

$$= 1 - p$$

*G. spekei*,  $N = 15$ :

rodiče 19 jedinců z 1979 jako referenční generace

→  $N_{eV} = 20,1$

( $N_{eF}$  pro tuto populaci byla  $48,1 \approx 239 \% N_{eV}$ )

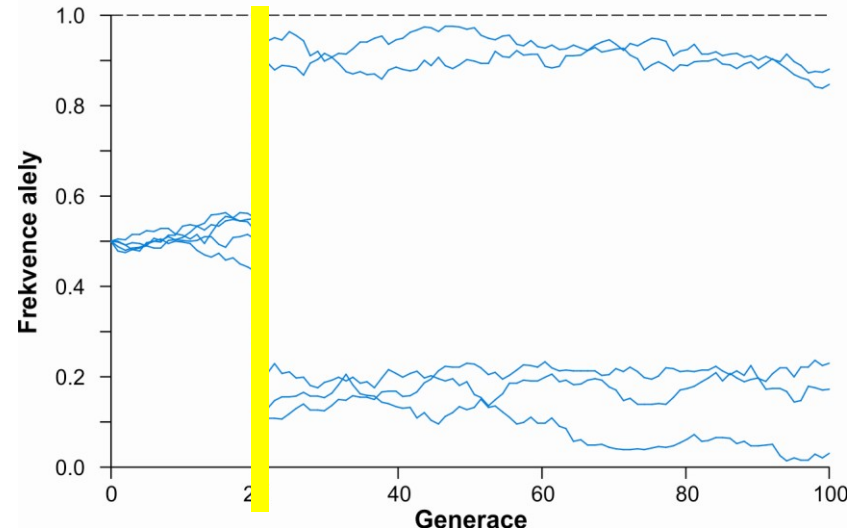


## Srovnání $N_{eF}$ a $N_{eV}$ :

$N_{eF}$  je citlivá k počtu rodičů

$N_{eV}$  je citlivá k počtu potomků

Př.: v 21. generaci snížení  
z  $N = 1000$  na  $N = 4$  jedince  
zvýšení opět na  $N = 1000$  ve 22. generaci:



$N_{eF}$ : gamety náhodně vybrány z 1000 rodičů  $\rightarrow$  4 jedinci v 21. generaci  
pravděpodobnost, že dvě vybrané gamety od stejného rodiče =  $1/1000$

pravděpodobnost, že se tyto gamety účastnily samooplození a jsou

$IBD = 1/2 \Rightarrow$  prům.  $F$  ve 21. generaci =  $1/(2N) = 1/2000$

$\Rightarrow \underline{N_{eF}(21) = 1000}$  (při  $N = 4$ )!

$N_{eV}$ : vybráno pouze 8 gamet  $\Rightarrow p(21) = pq/8 \Rightarrow \underline{N_{eV}(21) = 4}$  (při  $N = 4$ )

## Inbreeding, drift a ztráta alelické diverzity:

V mnoha reálných, početně omezených populacích sice existuje negativní korelace mezi inbreedingem a variabilitou ...

Ale ani děmový, ani rodokmenový inbreeding nemá přímý dopad na frekvence alel  $\Rightarrow$  sám o sobě ani nezrychluje, ani nezpomaluje ztrátu genetické variability

$\Rightarrow$  drift, ne inbreeding je skutečnou příčinou ztráty genetické variability

$\Rightarrow$  inbreeding nezpůsobuje ztrátu alelické diverzity v malých populacích, její příčinou je samotné omezení velikosti populace



## $N_e$ a rozdíly v počtu potomků (rozpor s Poissonovým rozdělením):

$k$  = počet potomků (*family size*)

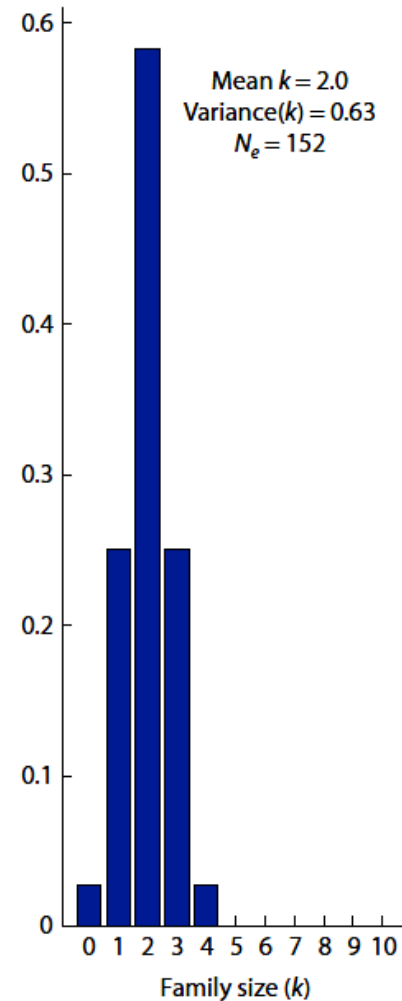
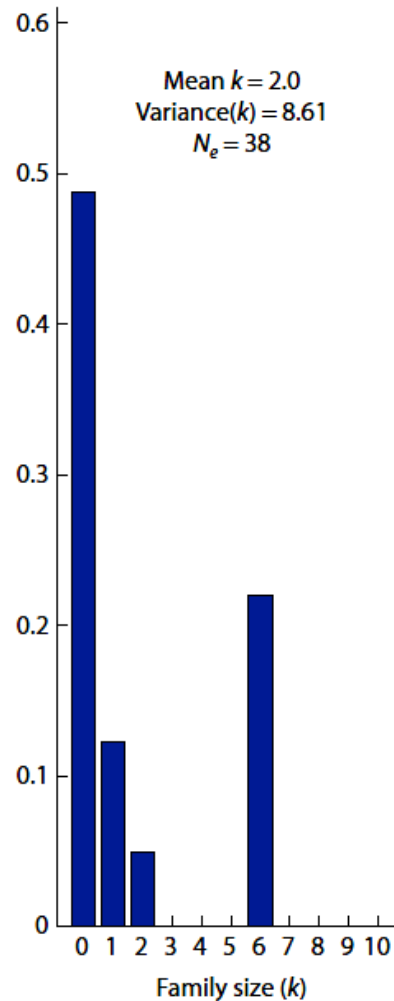
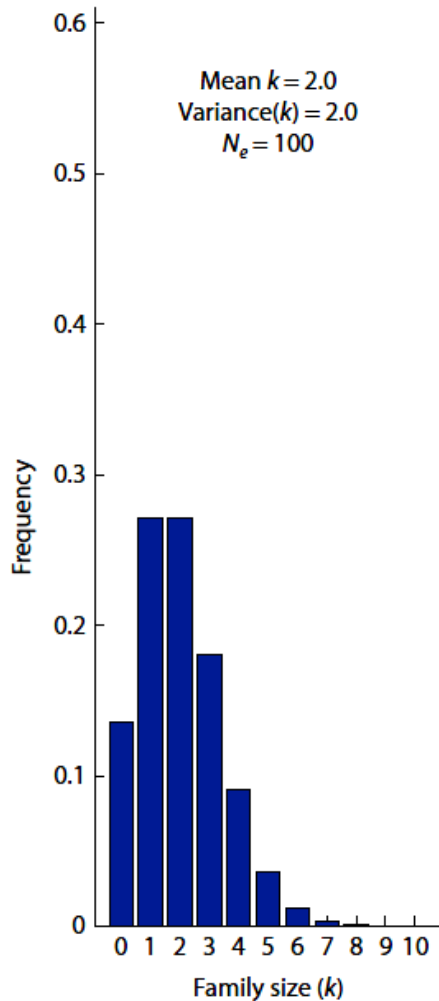
$N_e$  závisí na rozptylu  $k$

$$N_e = \frac{4N_{t-1}}{\text{Var}(k) + k^2 - k}$$

velikost populace  
rodičů

pro  $k = 2$ :

$$N_e = \frac{4N - 4}{\text{Var}(k) + 2}$$



Předpoklad WF modelu:  
 $\text{Var}(k) = \bar{k}$

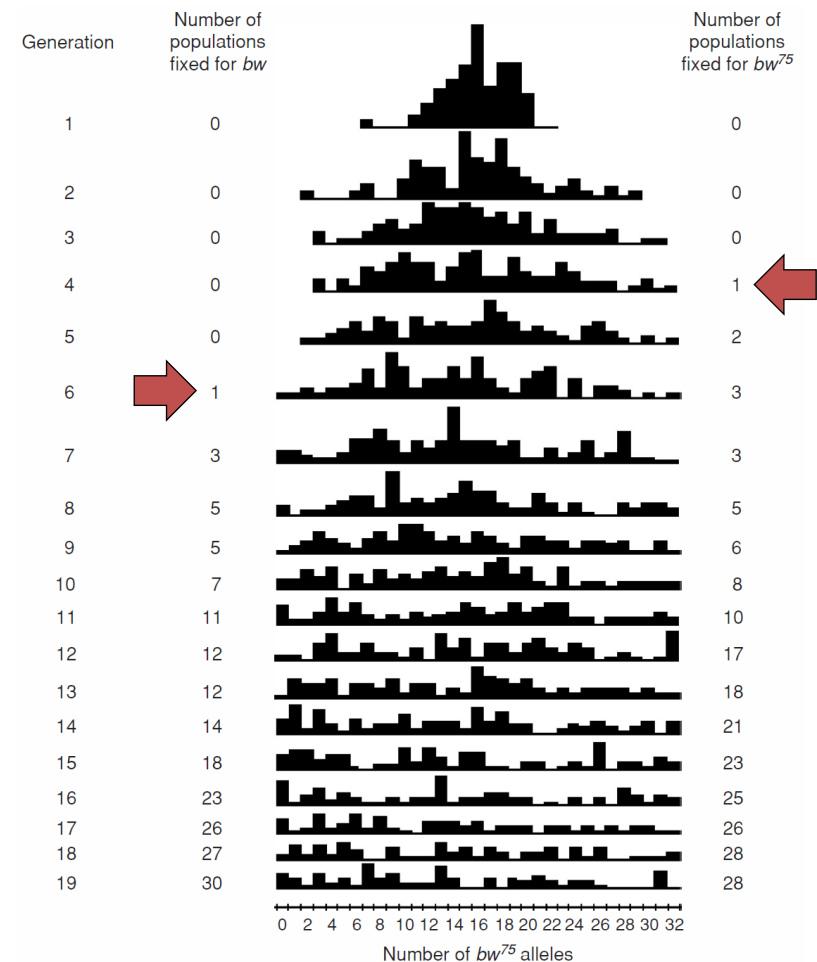
Pokud  $\text{Var}(k) > \bar{k}$ ,  
 $N_e < N$

Pokud  $\text{Var}(k) < \bar{k}$ ,  
 $N_e > N!$

V Buriho experimentu  $N$  v každé generaci = 8 samců a 8 samic ( $N = 16$ )  
 × jednotlivé populace dospívaly k fixaci/extinkci alel dříve než předpoklad,  
 což odpovídá menší populaci:  $N = 10-11$

důvodem např. menší počet samců, kteří se úspěšně reprodukovali nebo  
 vyšší rozptyl potomstva (tj.  $N_e < N$ )

např. pokud reprodukce jen 6 samců  
 a 8 samic a  $\text{Var}(k) = 3,5 \Rightarrow N_e \approx 10$



## Každý genetický znak vyžaduje vlastní $N_e$

Pro geny na autozomech, pohlavních chromozomech a mtDNA existuje odlišná efektivní velikost populace:

autozomy:  $N_e$  nebo  $4N_e$

X, Z:  $\frac{3}{4}N_e$  nebo  $3N_e$

Y, W, mtDNA:  $\frac{1}{4}N_e$  nebo  $1N_e$

# Vliv kolísání populační velikosti:

efektivní velikost lze aproximovat jako harmonický průměr  $\Rightarrow$  velký vliv malých  $N$ !!

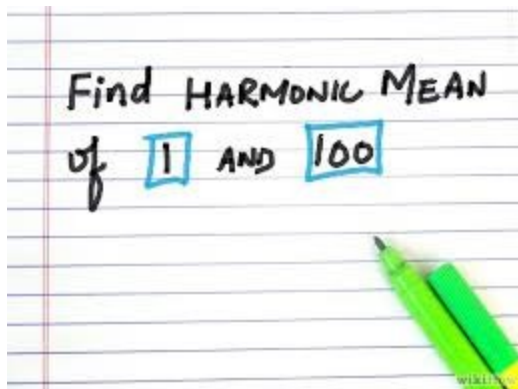
$$N_{eV} = \frac{t}{\frac{1}{N_1} + \frac{1}{N_2} + \dots + \frac{1}{N_t}}$$

harmonický průměr

$$N_{eF} = \frac{t}{\frac{1}{N_0} + \frac{1}{N_1} + \dots + \frac{1}{N_{t-1}}}$$

na rozdíl od  $N_{eV}$  závislost na rodičovské generaci

průměr mnohem blíží nižší hodnotě!



$$H = \frac{2}{\frac{1}{a_1} + \frac{1}{a_2}}$$
$$= \frac{2}{\frac{1}{1} + \frac{1}{100}}$$

$$= \frac{2}{1.01}$$
$$= 1.9802$$

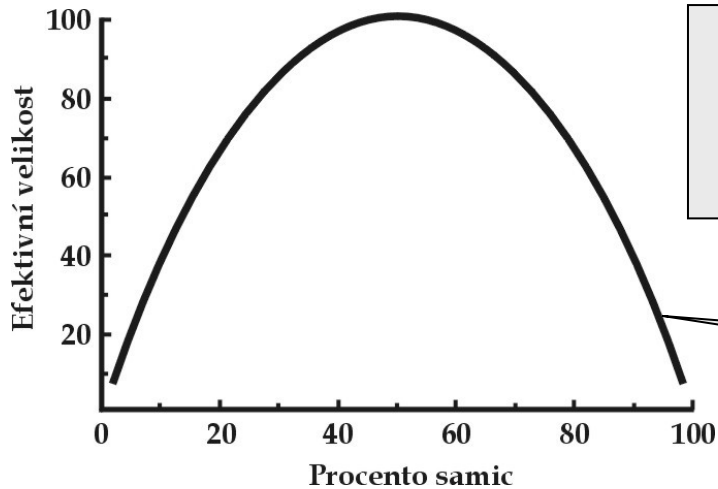
Harmonic Mean  
of 1 and 100

## Vliv vychýleného poměru pohlaví:

předchozí výpočty a aproximace předpokládaly stejný počet samců a samic přispívajících svými geny do dalších generací

$N_m$  = počet rozmnožujících se samců,  $N_f$  = počet samic

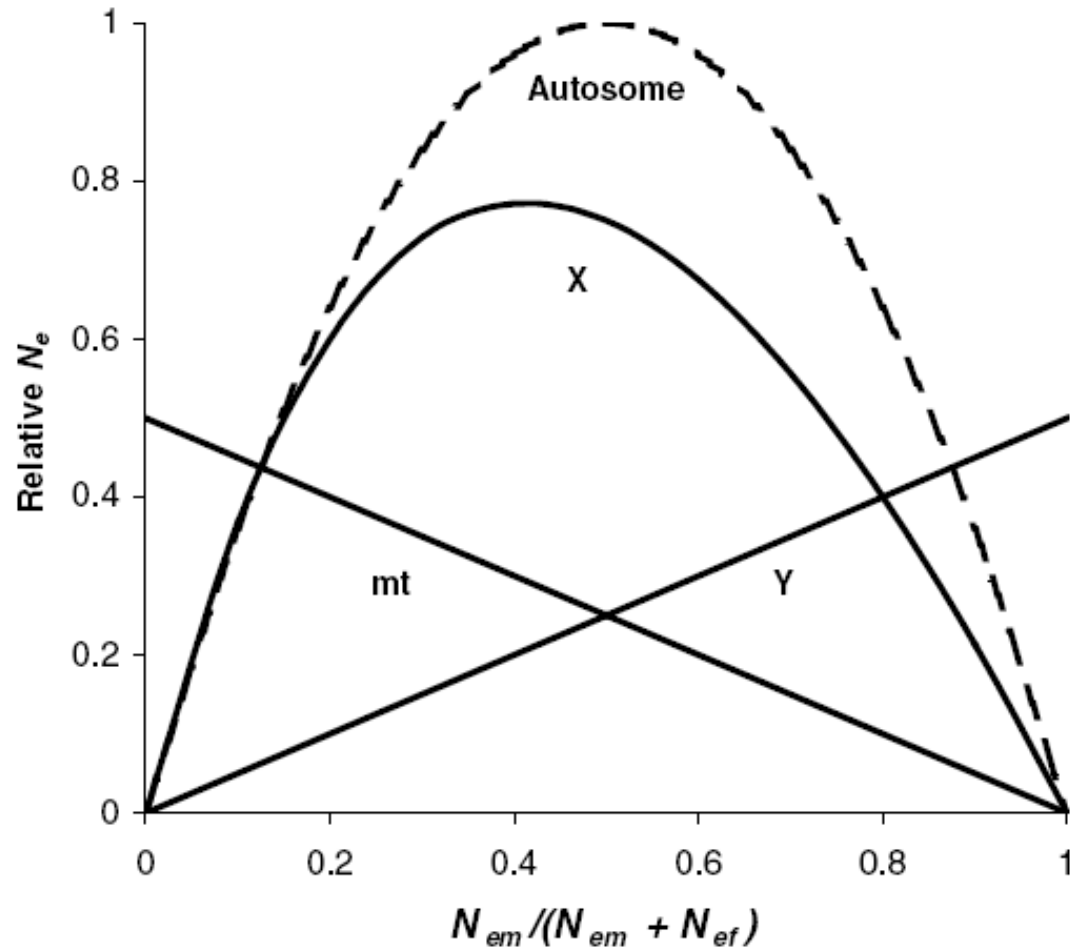
$$N_e = \frac{4N_m N_f}{n_m + n_f} \rightarrow N_e = \frac{4N_f}{n_m + n_f} \approx 4$$



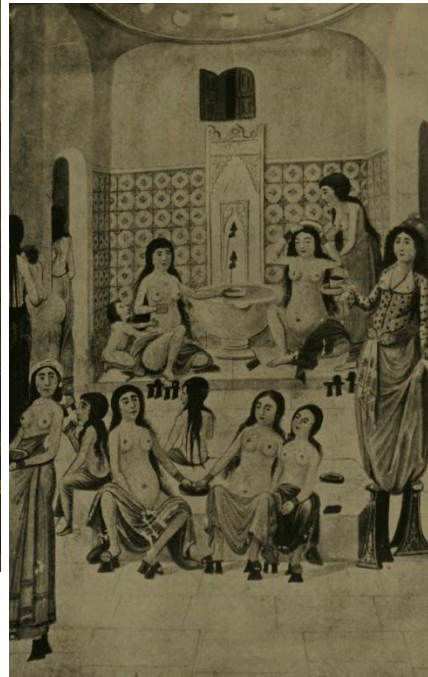
z uvedeného vztahu vyplývá, že pokud se v populaci bude rozmnožovat pouze jeden samec, bude  $N_e \approx 4$  bez ohledu na celkový počet jedinců

čím větší odchylky od vyrovnaného poměru pohlaví, tím nižší  $N_e$

vliv poměru pohlaví na  $N_e$  odlišný pro různé genetické znaky:



Př.: rypouš sloní (*Mirounga leonina*):  
v harému poměr pohlaví 1:40  
ale efektivní poměr 1:4–5 díky nevěrám a  
krátké době dominance samce (1–2 roky)





## Vliv selekce:

Jestliže na gen působí selekce, je rozptyl v počtu potomků mezi jedinci v populaci vysoký (jedinci s výhodnou alelou zanechají více potomstva)

⇒  $N_e$  pro tento gen nižší než pro gen selektivně neutrální

Platí i pro části genů!

## Pomocí koalescence můžeme definovat efektivní velikost populace $N_e$

$N_e$  = velikost idealizované populace, která má stejný koalescenční čas jako skutečná populace

Takto definovaná  $N_e$  součástí softwarů založených na koalescenci

Např. generační čas vyjádřen v jednotkách  $4N_e\mu$ , kde  $N_e$  = koalescenční efektivní velikost populace

Pozor: také v tomto případě rozdíly mezi geny:

oblasti s nízkou rekombinací → nízká  $N_e$

oblasti s vysokou rekombinací → vysoká  $N_e$