

Kvantifikace síly důkazního materiálu

doc. Mgr. Jiří Drábek, PhD.
Laboratoř experimentální medicíny,
Ústav molekulární a translační medicíny,
LF UP a Fakultní nemocnice Olomouc



Cíl lekce

- Po této lekci byste měli:
 - znát Bayesovu [čti: bezzovu] větu, rozumět jí a vědět, jak ji aplikovat u forenzních důkazů
 - zasadit důkaz na stupeň v hierarchii důkazů
 - uvědomovat si nedosažitelnost 100% jistoty
 - uvědomovat si omezení selského rozumu
 - umět vypočítat věrohodnostní poměr pro ztotožnění a jednoduché rodokmeny u testování příbuznosti.

Osnova přednášky

- Wasonův test, mamografické vyšetření
- Zákony pravděpodobnosti a Bayesova věta
- Rozložení Bayesovy věty na prvočinitele
- Hardy-Weinberg
- Ztotožnění
- Určování otcovství, odvození na rodokmenech

Ideální forenzně využitelný důkaz

- Má znaky, které jsou unikátní pro individuální osobu
- Tyto znaky se nemění v čase
- Znaky jsou jednoznačně určitelné na různých místech různými experty
- Umožňuje potvrdit přítomnost osoby na místě činu; vyskytuje se vždy, když jej potřebujeme pro potvrzení hypotézy a zároveň se nevyskytuje nikdy jindy
- Má jednoduché a levné zjištění hodnoty znaku

4

Typický forenzně využitelný důkaz

- Znak je přítomen, i když tvrzení není pravdivé (falešná pozitivita testu)
- Znak není přítomen, i když je tvrzení pravdivé (falešná negativita testu)
- Je pravděpodobnější, že se znak vyskytuje, pokud je tvrzení pravdivé

5

Omezení identifikačního důkazu

- Jakákoliv technika identifikace dokáže určit objekt s vysokou přesností za kontrolovaných podmínek. Místo činu v okamžiku činu nemá kontrolované podmínky.
- Subjektivitu je možno redukovat (objektivizovat), ale bez lidského zásahu a personifikace ohodnocení nejistoty to nejde.

6

Vlastnosti důkazu

- Nedá se použít samostatně
- Dá se kombinovat
- Pomáhá rozlišit mezi hypotézami
- Jeho síla závisí na hypotézách

Kousání nehtů a zneužívání

- 80 % zneužívaných dětí si kouše nehty
- 10 % nikdy nezneužívaných dětí si kouše nehty
- Dítě si kouše nehty. Známe pravděpodobnost kousání nehtů (důkazu) za dvou rozdílných scénářů. Neznáme pravděpodobnost scénářů. Je to problém inverzních pravděpodobností. Co s ním?

8

Pokud číselně nehodnotíme sílu důkazu, tak nevyužijeme sílu důkazu naplno

- **Podhodnotíme**
- Nadhodnotíme
 - CSI efekt

9

When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of science.

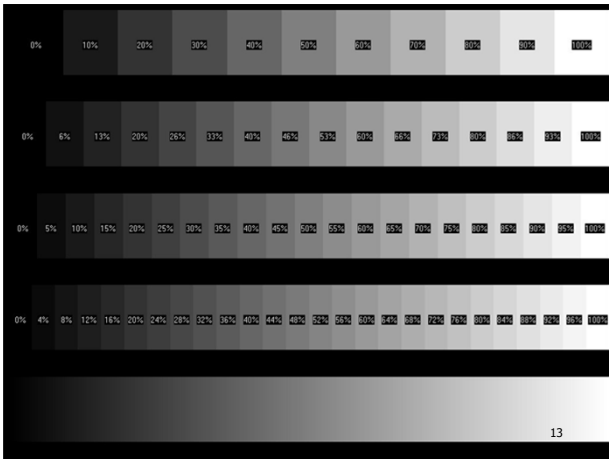
Lord Kelvin

Nelogický přístup soudu

- Shoda vs. vyloučení
 - 100% shoda, 100% vyloučení
 - Pokud není 100%, tak to není k ničemu
- Indicie
 - Pokud není 100%, tak je důkaz zamítnut, protože není „definitivní a jednoznačný“
 - Pokud není 100%, tak je to indicie a indicie jsou nedostačující

11

Černobílé



S intuicí a selským rozumem neuspějete

- Bayesovské uvažování
- Wasonův test obracení karet
- Kvantová mechanika

Wasonův test obracení karet

- Je obtížné rozhodnout se, co potřebujete vědět k rozhodnutí o pravdivosti logického problému
- Design testu: přednesení pravidla a otázka, zda je pravidlo porušeno

Wasonův test obracení karet

- **Pravidlo:** Pokud má karta „D“ na jedné straně, tak má „3“ na druhé straně.
- Předpoklad: každá karta má na jedné straně písmeno a na druhé straně číslo
- Falzifikace tvrzení – průkaz, že tvrzení je nesprávné
- Test: ukázány 4 karty : D, K, 3, 7
- Které dvě karty musíte otočit, abyste zjistili, jestli pravidlo platí?

D

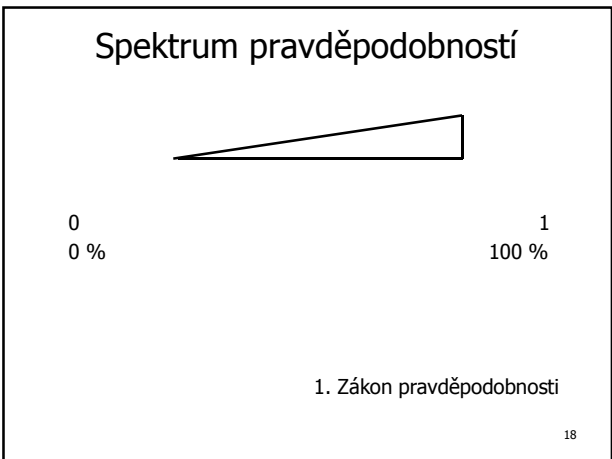
K

3

7

Wasonův test obracení karet

- Správná odpověď: D a 7
- Nejčastější špatná odpověď: D a 3
- Pokud otočíte kartu 7, čekáte D nebo K. Pokud bude D, falzifikovali jste pravidlo – pravidlo neplatí.
- Pokud otočíte kartu 3, čekáte D nebo K. Pokud bude D, potvrzujete pravidlo. Pokud bude K, tak nefalzifikujete pravidlo, nemůžete odpovědět.
- (Neexistence důkazu pro jev neznamena důkaz neexistence jevu).



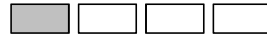
Další zákony pravděpodobnosti

- Pravděpodobnost dvou **současných**, na sobě nezávislých jevů počítáme jako součin pravděpodobností jednotlivých jevů („**a zároveň**“ implikuje „**krát**“)
- Pravděpodobnost dvou **alternativních** jevů počítáme jako součet pravděpodobností jednotlivých jevů („**bud' anebo**“ implikuje „**plus**“).

19

Pravděpodobnost a šance

- Pravděpodobnost 1/4 (25 %)



20

Pravděpodobnost a šance

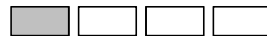
- Pravděpodobnost 1/4 (25 %)



21

Pravděpodobnost a šance

- Pravděpodobnost 1 děleno 4 odpovídá šanci 1 ku 3



22

Pravděpodobnost a šance

- Pravděpodobnost 1 děleno 4 odpovídá šanci 1 ku 3



23

Převod šance na pravděpodobnost

Šance x ku y
Pravděpodobnost = $x/(x+y)$

Převod pravděpodobnosti na šanci

Pravděpodobnost v rozsahu 0 až 1
Šance = $P/(1-P)$

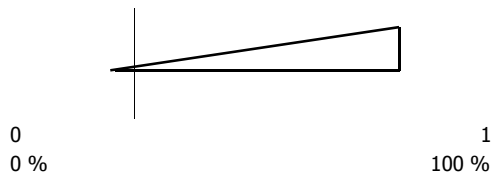
24

Logický přístup soudu - rozdělení úloh a použití Bayesovy věty

- Znalec
 - Nabídka cena:výkon
 - Výpočet věrohodnostního poměru
 - Doporučení alternativních hypotéz
 - Návod k interpretaci
- Soudce
 - Stanovení alternativních hypotéz
 - Určení apriorní pravděpodobnosti viny
 - Přijetí/nepřijetí důkazu
 - Pokud je důkaz přijat, tak stanovení hranice, kdy je vina prakticky prokázána
- Obhájce
 - Nabídka alternativních hypotéz

25

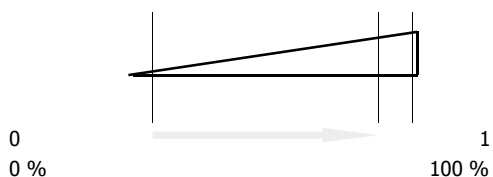
Počáteční ohodnocení pravděpodobnosti



Má provádět soud

26

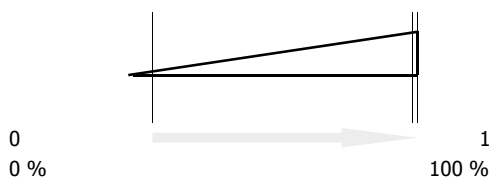
Úprava ohodnocení pravděpodobnosti



Znalec má pravděpodobnostní očekávání soudu upravit na základě odborného vyšetření.
Soud stanoví prahovou hodnotu.
Zde prahová hodnota nedosažena.

27

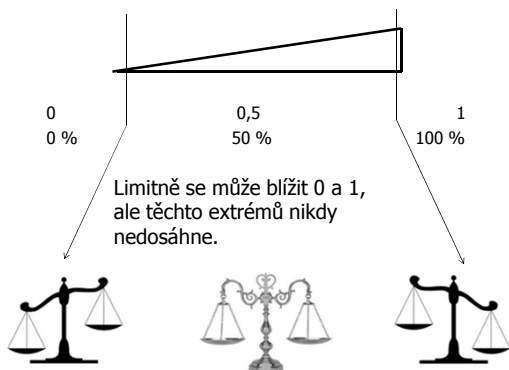
Úprava ohodnocení pravděpodobnosti



Prahová hodnota překročena

28

Spektrum pravděpodobností



Bayesova věta

- Umožňuje (kvantitativně) kombinovat důkazy s obecnou znalostí o případu
- Používá se v bankovníctví, pro filtrování spamu v emailu, pro studium evoluce, plánování klinických zkoušek...
- Není to novinka (1763). Vyžaduje práci s inverzní pravděpodobností, což neodpovídá klasickému (frekvencistickému) statistickému přístupu ani intuici.

30

Historie bayesovského přístupu

- Thomas Bayes (1763): jak změnit naši důvěru v hypotézu ve světle nových důkazů
- Essen-Moller (1938): pravděpodobnost otcovství z krevní zkoušky
- Alan Turing (1940): úprava pravděpodobnosti pro prolomení kódu ve 2. světové válce
- Jack Good (1950): klasické dílo „Pravděpodobnost a vážení důkazů“
- Dennis Lindley (1970): „Jak pochopit nejistotu“ – aplikace LR ve forenzní vědě
- Buckleton, Evett, Weir, Curran (1990) interpretace DNA důkazů u směsí

31

Bayesovské uvažování v praxi: medicínská informace

- U nás je mamografický screening pro ženy po 45 letech každé dva roky zdarma
- Vyšetření trvá 10 minut
- Včasně zachycenou rakovinu prsu lze úplně vyléčit
- Později zachycená rakovina může skončit ablací prsu
- Nezachycená rakovina může skončit smrtí

Bayesovské uvažování v praxi (částečně smyšlená data)

Premisy

- **1%** ze všech pětáctýřicetiletých žen, které chodí na rutinní preventivní skrining, má rakovinu prsu
- **80%** žen s rakovinou prsu má pozitivní mamografický nález (falešná negativita **20%**)
- **9,6%** žen bez rakoviny prsu má pozitivní mamografický nález (falešná pozitivita **9,6%**)

Nález

- 45 letá žena má pozitivní mamografický nález při preventivním skriningu

Otázka

- Jaká je pravděpodobnost, že tato žena má rakovinu prsu?

Tabulkově pro 100 000 žen

	S rakovinou	Zdravé	
Pozitivní mamografický nález			
Negativní mamografický nález			

Tabulkově pro 100 000 žen

- **1%** má rakovinu prsu

	S rakovinou	Zdravé	
Pozitivní mamografický nález			
Negativní mamografický nález			
	1000	99 000	100 000

Tabulkově pro 100 000 žen

- **80%** žen s rakovinou prsu má pozitivní mamografický nález

	S rakovinou	Zdravé	
Pozitivní mamografický nález	800		
Negativní mamografický nález	200		
	1 000	99 000	100 000

Tabulkově pro 100 000 žen			
• 9,6% žen bez rakoviny prsu má pozitivní mamografický nález			
	S rakovinou	Zdravé	
Pozitivní mamografický nález	800	9 504	10 304
Negativní mamografický nález	200	89 496	89 696
	1 000	99 000	100 000

Tabulkově pro 100 000 žen			
	S rakovinou	Zdravé	
Pozitivní mamografický nález	800	9 504	10 304
Negativní mamografický nález	200	89 496	89 696
	1 000	99 000	100 000

$$800/10304 = 0,0776$$

- Pravdivě (správně) pozitivní, true positive

Správný postup

- Použít všechny tři známé pravděpodobnosti:
 - % žen s rakovinou
 - % žen bez rakoviny a s pozitivním nálezem
 - % žen s rakovinou a s pozitivním nálezem
- (čili započítat prevalenci, senzitivitu a specificitu).

Odvození správné odpovědi

- 1 000 žen z 100 000 má rakovinu prsu (1%)
 - 800 žen z těchto 1 000 má pozitivní nález (80%)
- 99 000 žen z 100 000 nemá rakovinu prsu (99%)
 - 9 504 z těchto 99 000 zdravých žen (9,6%) má pozitivní nález
- Celkem je $9504+800=10\,304$ pozitivních nálezů
- Celkem je 800 žen s rakovinou
- Procento žen s rakovinou po pozitivní mamografii je $800/(10\,304) = 0,0776397$ (**7,8%**).

Nejčastější chyba

- Ignorování počtu žen s rakovinou (např. nesprávný výsledek 80%)

Symbolika pro kvantifikátory

P	pravděpodobnost
E	důkaz (evidence, zde výsledek typizace)
H	hypotéza
	za předpokladu
m	počet genotypů
G _a	frekvence osob s genotypem „a“
N	počet alel
o	počet osob
p _i	frekvence alely „i“
r	počet typizačních systémů
μ	četnost mutace

Bayesův teorém

- Úprava pravděpodobnostního očekávání ve světle nových důkazů
- Pravidlo pro rozhodování mezi více alternativami
 - Logické, konzistentní, univerzální

Bayesův teorém v šancovém tvaru

Aposteriorní šance = apriorní šance * věrohodnostní poměr

$$\text{Apo} = \text{apri} * \text{věrohodnostní poměr}$$

45

Bayesova věta v pravděpodobnostní formě

$$PO = \frac{PO_{\text{výchozí}} * LR}{PO_{\text{výchozí}} * LR + 1 - PO_{\text{výchozí}}}$$

U určování otcovství se při „neutrální“ 50% výchozí pravděpodobnosti zjednodušuje na

$$PO_{50} = \frac{LR}{LR + 1}$$

Bayesův teorém v šancovém tvaru

Aposteriorní šance = apriorní šance * věrohodnostní poměr

$$\text{Apo} = \text{apri} * \text{věrohodnostní poměr}$$

47

Apriorní šance

- Měřítka nejistoty ohledně viny nařčeného předtím, než je provedeno znalecké dokazování.
- Na základě známých faktů o případu, k jejichž zhodnocení nejsou potřebné expertní znalosti.
- Může být odvozeno z frekvence výskytu jevu v populaci (prevalence), z encyklopedických údajů nebo z životních zkušeností.
- Má určovat soud.

48

Příklady apriorní šance

- Osoba nakažena virem HIV
– 1 ku 3 000
- Nepovolenou rychlostí jedoucí řidič požil alkohol
– 1 ku 5
- Člověk má na sobě skleněné střípky
 - 1 ku 1 pro sklenáře
 - 1 ku 3 pro bezdomovce
 - 1 ku 50 pro kutila
 - 1 ku 5 000 pro kancelářskou myš

49

Jednotná 50% pravděpodobnost pro paternity a 1% pro ztotožnění?

50

Bayesův teorém v podobě šancí

$A_{po} = a_{pri} * \text{věrohodnostní poměr}$

51

Věrohodnostní poměr
aktualizuje pravděpodobnost

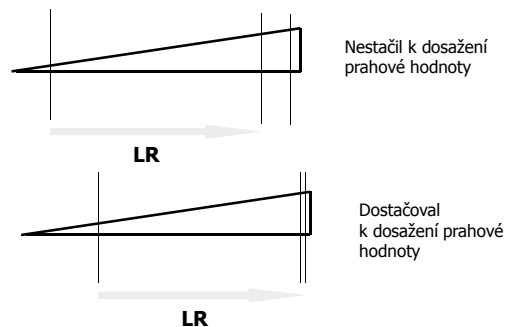
Co je to věrohodnostní poměr?

- Jediné číslo, které sumarizuje podporu jedné hypotézy započtením pravděpodobnosti důkazu za předpokladu jedné hypotézy a pravděpodobnosti stejného důkazu za předpokladu druhé hypotézy.

$$LR = \frac{P(E|H1)}{P(E|H2)}$$

53

Věrohodnostní poměr, likelihood ratio (LR)



54

Hodnoty LR

- LR = 1 důkaz je irelevantní, nepodporuje ani jednu hypotézu
- LR > 1 důkaz podporuje naši hypotézu
- LR < 1 důkaz podporuje alternativní hypotézu

Kdy se ukáže výhoda bayesovského přístupu a věrohodnostního poměru?

- Tři alely
- Nulová alela
- Mutace
- Laboratorní chyba
- Komplikovaný rodokmen
- Směsi se zvýšenými stochastickými efekty
- Více typů důkazů

Převod čísel na slovní vyjádření

- | • LR rozsah | Slovně |
|----------------|-------------------------|
| • <0,001 | velmi silný důkaz proti |
| • 0,001 ÷ 0,01 | silný důkaz důkaz proti |
| • 0,01 ÷ 0,1 | středně silný proti |
| • 0,1 ÷ 0,99 | slabý důkaz proti |
| • 1 ÷ 10 | slabý důkaz pro |
| • 10 ÷ 100 | středně silný důkaz pro |
| • 100 ÷ 1000 | silný důkaz pro |
| • >1000 | velmi silný důkaz pro |

57

Nebezpečí hraničního efektu

Kousání nehtů a zneužívání – zhodnocení věrohodnostním poměrem

- 80 % zneužívaných dětí si kouše nehty
 - 10 % nikdy nezneužívaných dětí si kouše nehty
- Věrohodnostní poměr

$$LR = \frac{P(E|H1)}{P(E|H2)} = \frac{0,8}{0,1} = 8$$

59

Mamografický příklad pomocí šancové podoby Bayesovy věty

1% prevalence odpovídá kurzu 1 ku 99
80% správně pozitivních testů (20% falešná negativita)
9,6% falešně pozitivních testů

Aposteriorní šance = apriorní šance * věrohodnostní poměr
Aposteriorní šance = 1/99 * 0,8/0,096 = 0,084175 ku 1

Převod šance na pravděpodobnost = šance/(šance + 1) = Aposteriorní šance = 0,084175/1,084175 = 0,0776396 ~ 7,8%

<http://yudkowsky.net/bayes/bayes.html>

60

Hypotézy, scénáře, verze, teze, tvrzení, varianty, vysvětlení

- Vždy musíme srovnávat nejméně dvě navzájem se vylučující hypotézy.
- PŘ.: Na místě činu byla nalezena krevní skvrna, jejíž DNA profil je shodný s profilem podezřelého.
- H1: definuje žalobce
 - Podezřelý zanechal na místě činu krevní skvrnu
- H2: definuje obhájce
 - Někdo jiný, neznámý, zanechal na místě činu krevní skvrnu. Podezřelý má náhodou shodný profil jako neznámý pachatel.

61

Alternativní hypotéza

- Podezřelého srovnáváme s náhodným mužem z populace jen tehdy, pokud není důkaz, který by zúžil obecnou populaci, a nejedná se o subpopulaci podezřelých.
- Alternativní hypotézu dodává obhájce; může se jednat o sloučení několika hypotéz.
- Hypotézy nemusí být vyčerpávající (nemusí zahrnovat všechna možná vysvětlení).
- Alternativní hypotézy musejí být představitelné a s reálnou apriorní pravděpodobností (na základě důkazů a obecných znalostí).

62

Hierarchie hypotéz

- Jeho biologický materiál je na střece nože
- Vs.
- Biologický materiál na střece noži patří neznámému muži

63

...hierarchie hypotéz

- Držel nůž v ruce
 - Jeho biologický materiál je na střece nože
- Nutno vzít v úvahu vliv času, přilnavosti, trvanlivosti, přenosu

64

...hierarchie hypotéz

- Nožem vraždil
 - Držel nůž v ruce
 - Jeho biologický materiál je na střece nože
- Nutno vzít v úvahu kinetiku pohybu, sílu, koncept přiměřené obrany

65

...hierarchie hypotéz

- Nožem vraždil
 - Držel nůž v ruce
 - Jeho biologický materiál je na střece nože
- Level/Úroveň**
- | | |
|------------|----------------|
| – Offence | – Zločinu |
| • Activity | • Činnosti |
| – Source | – Zdroje stopy |

66

...hierarchie hypotéz

Úroveň

- Zločinu
 - Činnosti
 - Zdroje stopy



**Hodnota důkazu
Informace
Požadovaná
expertíza**

67

Čtvrtá úroveň hypotéz

Úroveň

- Zločinu
 - Činnosti
 - Zdroje stopy
 - Zdroje mikrostopy

Čtvrtá úroveň tvrzení přidána kvůli zvyšující se citivosti DNA profilování.

Jsou to hypotézy:

DNA profil patří podezřelému

Versus

DNA profil patří neznámému muži z populace

68

Vliv soutěžících hypotéz na sílu důkazu

- Ignorování apriorních P a relevantních hypotéz je chybné. Pravděpodobnost je vždy podmíněná. Vždy záleží na okolnostech.
- DNA důkaz prokáže přítomnost osoby na místě činu
 - K ničemu, pokud se jedná o „domácí“ osobu
- DNA důkaz prokáže semeno na kalhotkách
 - K ničemu, pokud je alternativní hypotézou, že sex byl konsenzuální
- Nařčený má 5 znaků biologického otce z 5 testovaných:
 - Pokud nařčený tvrdí: „To ne já, to některý z mých bratrů!“, síla důkazu se snižuje.

69

Výpočet síly DNA důkazu

Předpoklady

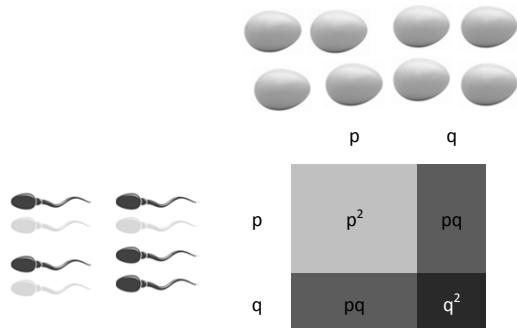
- Platnost Mendelova zákona segregace a volné kombinovatelnosti vloh
- Hardy Weinberg Equilibrium (HWE)
- Geny nesmí být ve vazbě, aby se daly násobit jednotlivé věrohodnostní poměry při výpočtu celkového věrohodnostního poměru (multiplikativnost).

Hardy Weinbergova rovnováha (HWE)

- Vzorec popisující populaci v rovnováze
- Za splnění určitých podmínek (panmixie, velká populace bez migrace, bez genetického driftu a selekčního tlaku) se po jedné generaci fixují genotypové frekvence jako funkce alelických frekvencí.
- Když známe jedno, dopočteme druhé.

72

Punnetův čtverec pro H-W



Hardy Weinbergova rovnováha

Počet alel	Názvy alel	Frekvence alel	Frekvence genotypů <i>Genotypy</i>
2	P, Q	p, q	$(p+q)^2 = p^2 + 2pq + q^2$ <i>PP+PQ+QQ</i>
3	P, Q, R	p, q, r	$(p+q+r)^2 = p^2 + 2pq + 2qr + 2pr + q^2 + r^2$ <i>PP+PQ+QR+PR+QQ+RR</i>
4	P, Q, R, S	p, q, r, s	$(p+q+r+s)^2 = p^2 + 2pq + 2qr + 2pr + 2ps + 2rs + 2qs + q^2 + r^2 + s^2$ <i>PP+PQ+QR+PR+PS+RS+QS+QQ+RR+SS</i>

Panmixie - neasortativní páření

Genetický drift – efekt hrdla láhve

76

Podmínky pro HWE nejsou splněny?

- Matematický model je ideál
- Přestože je ideál nedosažitelný, může sloužit jako aproximace.
- Pokud aproximace nedostačuje, je možno použít korekce (théta korekce pro subpopulace a podobně)

77

Využití HWE

- Známe frekvenci $p (=0,01)$ alely P a frekvenci $q (=0,03)$ alely Q
- Heterozygot PQ má populační frekvenci $2 * p * q = 2 * 0,01 * 0,03 = 0,0006$
- Průměrně 6 lidí z 10000 bude mít náš genotyp PQ
- Hodnotu markanty určuje její vzácnost.

Pokud se jedná o ztotožnění

Biologická stopa



79

... ztotožnění

Biologická stopa $\xrightarrow{\text{test}}$ Hrubá data



80

... ztotožnění

Biologická stopa $\xrightarrow{\text{test}}$ Hrubá data $\xrightarrow{\text{odečet}}$ Genotyp stopy S
11, 13



81

... ztotožnění

Biologická stopa $\xrightarrow{\text{test}}$ Hrubá data $\xrightarrow{\text{odečet}}$ Genotyp stopy S
11, 13

\Downarrow Srovnání

Genotyp podezřelého P
11, 13



82

... ztotožnění

Biologická stopa $\xrightarrow{\text{test}}$ Hrubá data $\xrightarrow{\text{odečet}}$ Genotyp stopy S
11, 13

\Downarrow Srovnání

Genotyp podezřelého P
11, 13

$P(\text{identifikace}) =$
 $= P(\text{DNA profil stopy a podezřelého jsou shodný} |$
 $\text{pocházejí ze stejného zdroje}) = 1$

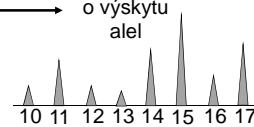


83

Pokud se jedná o náhodnou shodu

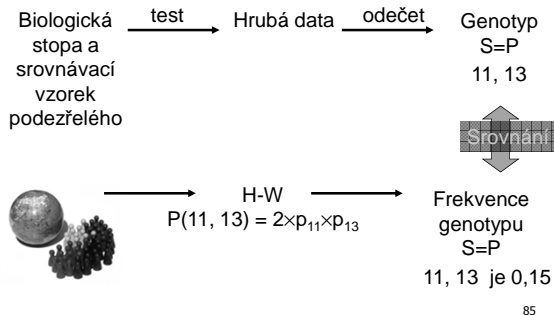
Biologická stopa a srovnávací vzorek podezřelého $\xrightarrow{\text{test}}$ Hrubá data $\xrightarrow{\text{odečet}}$ Genotyp S=P
11, 13

Populační data o výskytu alel



84

Pokud se jedná o náhodnou shodu



85

Pokud se jedná o náhodnou shodu

P(koincidence) =
= P(DNA profil stopy a podezřelého jsou shodně] pocházejí z rozdílných zdrojů) = 0,15

Věrohodnostní poměr u ztotožnění

alternativní hypotéza:
Stopu zanechal někdo jiný

Za předpokladu, že...
Pokud...

Hypotézy stavíme do kontrastu = $\frac{P(\text{DNA profilu} \mid \text{identifikační hypotéza})}{P(\text{DNA profilu} \mid \text{alternativní hypotéza})}$

Pravděpodobnost DNA profilu je (milionkrát) větší, pokud podezřelý zanechal stopu než pokud někdo jiný zanechal stopu.

87

...věrohodnostní poměr u ztotožnění

$\frac{P(\text{DNA profil stopy a podezřelého jsou shodně] pocházejí ze stejného zdroje})}{P(\text{DNA profil stopy a podezřelého jsou shodně] pocházejí z rozdílných zdrojů})} = \frac{1}{0,15} = 6,7$

LR=1/RMNE

RMNE=Random Man Not Excluded

88

Komplikace u kriminálních případů

- Degradovaná DNA poskytující jen částečný profil, použití LCN (low copy number) modifikací metody
 - alelický dropout
 - alelický dropin
- Smíšené vzorky
 - komplikovanější statistika
 - někdy nelze získat výsledek

Jak ovlivní etnikum podezřelého výběr populační databáze?

- Populační databáze se týká možných podezřelých
- Není důležité etnikum podezřelého, ale etnikum pachatele zločinu, pokud jím není náš podezřelý
- Íránský lékař v USA je podezřelý, že znásilnil pacientku. Jeho DNA profil se shoduje s částečným profilem získaném ze stop po znásilnění. Obhájce tvrdí, že pro výpočet síly důkazu není správné použít populační frekvence z celoamerické databáze, ale že se mají použít frekvence z íránské vesnice, odkud lékař pochází, a kde je vysoká úroveň konsagvinity.
- Má advokát pravdu?

90

Populační databáze

- Má odpovídat té populaci, ze které se rekrutuje pachatel, pokud pachatelem není podezřelý.

Testování otcovství

- Zákon o rodině č.94/163 Sb., §54 odst. 2: za otce dítěte se považuje muž, který s matkou souložil v době, od které neprošlo do narození dítěte méně než 180 a více než 300 dnů, pokud jeho otcovství závažné okolnosti nevylučují (doba rozhodná)
- Při sporném otcovství se uplatňují důkazy
 - gynekologické (doba početí)
 - andrologické (neplodnost z důvodu *impotentia coeundi* a *impotentia generandi*)
 - genetické (genotyp na úrovni molgen nebo fenotyp krve, vnějších znaků).

Princip testování otcovství

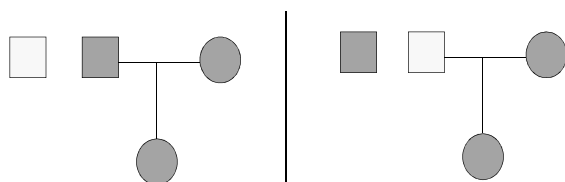
- Obvykle trio Matka, Dítě, Žalovaný muž (zvaný také testovaný, nařčený, pravděpodobný otec)
- Dítě má dvě alely od každého autozomálního genu

Mater semper certa est

Ale:

- IVF (*in vitro* fertilizace), „foster mother“ dítě donosí, další žena dítě adoptuje
- Čínské, české, německé porodnice
- Babyboxy
- Chiméry

Je nařčený muž biologickým otcem dítěte?

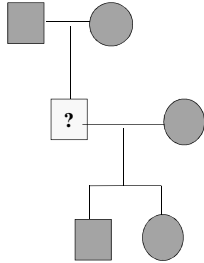


oranžově testované osoby

Hypotézy paternity

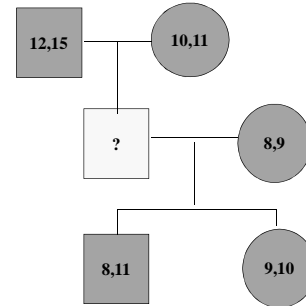
- H1: nařčený muž je biologickým otcem
- H2: jiný neznámý muž je biologickým otcem a náhodou (koincidencí) má nařčený stejné znaky jako biologický otec
- H3: nařčený muž č.2 je biologickým otcem
- H4: blízký příbuzný nařčeného je biologickým otcem

Kontrolní otázka: jaké hypotézy otcovství mohu zkoumat u tohoto případu?



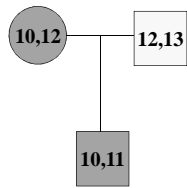
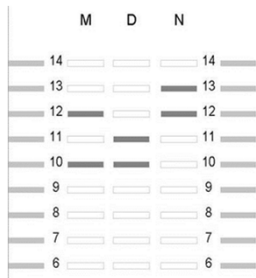
- otcem je neznámý muž
- otcem jsou dva neznámí muži
- chlapec patří nařčenému
- holčička patří nařčenému

Kontrolní otázka: může být nařčený otcem obou dětí?

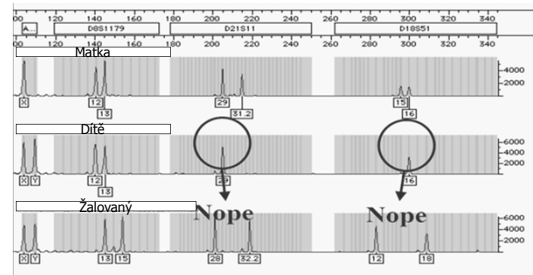


D16S539

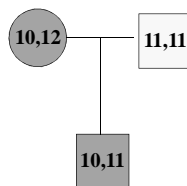
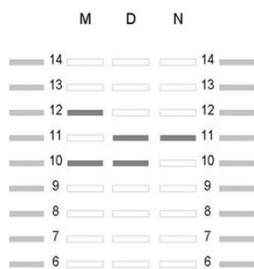
Žalovaný nemá alelu biologického otce



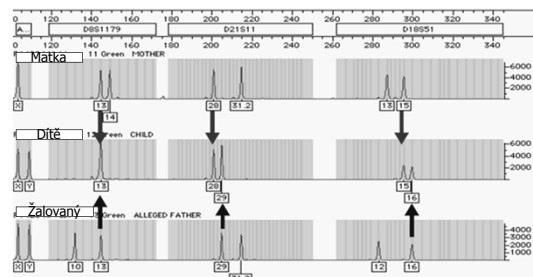
Žalovaný nemá alely biologického otce



Žalovaný nese alelu biologického otce



Žalovaný nese alely biologického otce



Způsoby vyjádření síly důkazu po typizaci u paternity

- **LR (Likelihood Ratio) = věrohodnostní poměr = PI (Paternity Index) = index otcovství**
- Kritické číslo = obrácená hodnota PI
- **W (Wahrscheinlichkeit) = PO = Pravděpodobnost otcovství**

Nedoporučované způsoby vyjádření síly důkazu po typizaci u paternity

- RMP, PM (Random Match Probability, Probability of Match) = Pravděpodobnost náhodné shody, Pravděpodobnost nevyloučení ne-otce.
- PE, PD (Probability of Exclusion, Power of Exclusion, Power of Discrimination) = Pravděpodobnost vyloučení, Diskriminační síla

Pravděpodobnost náhodné shody

- Random match probability (RMP), Random man not excluded (RMNE)
- Pravděpodobnost, že dva náhodně vybraní lidé z populace budou mít *stejný* genotyp (frekvence všech mužů v populaci, kteří mají stejnou alelu, jako biologický otec dítěte)
- Součin pravděpodobností genotypů jednotlivých lokusů, které nejsou ve vazbě

$$PM = \sum_{a=1}^m G_a^2$$

Pravděpodobnost vyloučení (PE)

- Pravděpodobnost vyloučení náhodného muže z populace (ne-otce) z otcovství za předpokladu daného genotypizačního výsledku matky a dítěte.
- Genotypizační výsledek žalovaného se nebere v úvahu.
- PE = frekvence takových mužů v populaci, kteří nemají alelu biologického otce (frekvence neprávem nařčených mužů v populaci, frekvence ne-otců).

Pro výpočet PE potřebujeme znát

$$p_A$$

p_A = populační frekvence té alely dítěte, kterou zdědilo po svém biologickém otci

p_A se liší mezi subpopulacemi. Vybíráme nejpříhodnější populaci nebo celkovou populaci (popřípadě modifikovanou koeficientem příbuznosti théta).

PE

$$PE = 1 - (p_A^2 + 2p_A(1-p_A)) =$$

$$= 1 - (p_A^2 + 2p_A - 2p_A^2) =$$

$$= 1 - (2p_A - p_A^2) = 1 - 2p_A + p_A^2 =$$

$$= (1-p_A)^2$$

PE

- Pokud má matka i dítě genotyp AB, tak nevyloučení muži jsou ti, kteří nesou alelu A nebo alelu B nebo obě alely. V takovém případě:

$$PE = (1-(p_A+p_B))^2$$

Kombinovaná pravděpodobnost vyloučení (CPE)

- Kombinovaná pravděpodobnost vyloučení falešně nařčených mužů vyšetřením N genetických systémů za předpokladu nezávislosti systémů
- $CPE = ((1-PE_1)*(1-PE_2)*(1-PE_3)*...*(1-PE_N))$
 $= \Pi(1-PE)$

Pravděpodobnost náhodné shody vyjádřená pomocí PE

$$RMP = p_A^2 + 2p_A(1-p_A) = 1-PE$$

$1-p_A$ frekvence všech ne-A alel dohromady

Pravděpodobnost otcovství

- Měřitko síly důvěry v hypotézu, že žalovaný je biologickým otcem
- Pravděpodobnost otcovství musí vycházet z apriorní pravděpodobnosti a musí být vypočtena na základě všech důkazů v rámci případu
- Negenetické důkazy pocházejí ze svědectví matky, žalovaného a dalších svědků
- Genetické důkazy pocházejí z genotypizace tria M-D-Ž.

Pokusy znehodnotit/nadhodnotit výsledky DNA profilování

- Strana obhajoby
 - nedodání vzorku
 - dodání špatného vzorku
 - ignorování genetických zákonů
 - ignorování dalších důkazů
 - psychologické triky, procenta × desetinná čísla, logické triky
- Strana obžaloby
 - přehození příčiny a následku (prosecutor's fallacy)

Logický trik obhajoby

Ignorování dalších důkazů
Vydávání pravděpodobnosti genotypu nařčeného za pravděpodobnost, že to není on.

Případ O.J.Simpson

„Důkaz by měl být zamítnut, protože lidmi s takovými krevními skupinami, jako má O. J. Simpson, se dá naplnit celý fotbalový stadion.“

O.J.Simpson

- Soud století
- Drama spravedlnosti
- Elementy podmanivého příběhu: bohatství, sláva, láska, žárlivost, zrada

O. J. Simpson

- Orenthal James „O. J.“ Simpson, s přezdívkou „džus“ je bývalý hráč amerického fotbalu a odsouzený kriminálník.
- V roce 1985 se dostal do Síně slávy amerického fotbalu a slavil úspěchy jako herec a sportovní komentátor
- V roce 1995 byl v trestním soudním procesu the People vs Simpson osvobozen z obvinění, že zavraždil svou bývalou manželku Nicole Brown Simpson a jejího přítele Ronalda Goldmana
- V roce 1997 byl v civilním soudním procesu shledán vinným. Jako trest měl zaplatit 33,5 milionů \$, což zatím neučinil.
- V roce 2007 byl zatčen v Las Vegas, Nevada a obviněn z loupeže a únosu.
- V roce 2008 byl odsouzen k 33 letům vězení s možností být propuštěn na podmínku nejdříve po 9 letech. Sedí v the Lovelock Correctional Center in Lovelock, Nevada.

Záměna podmínky a důsledku

- Klam obousměrnosti
 - Téměř všichni teroristé jsou muslimové
 - neznámá, že
 - Téměř všichni muslimové jsou teroristé
 - (prázdný sud nejlíc duní)
- Prosecutor 's fallacy

118

Záměna podmínky a důsledku

Tvrzení, že $P(\text{důkazu}|H) = P(H|\text{důkazu})$, což není pravda

Př.1: $P(\text{zvíře má čtyři nohy}|\text{zvíře je pes}) \neq P(\text{zvíře je pes}|\text{zvíře má čtyři nohy})$

... záměna podmínky a důsledku

$P(\text{DNA profil podezřelého a stopy z místa činu jsou shodné}|\text{podezřelý byl na místě činu})$

\neq

$P(\text{podezřelý byl na místě činu}|\text{DNA profil podezřelého a stopy z místa činu jsou shodné})$

120

Případ birminghamské šestky

„Je na 99% procent jisté, že tito muži zacházeli s výbušninami.“

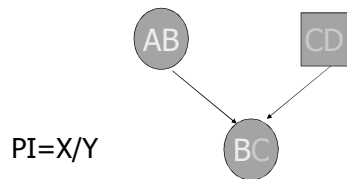
Logický trik obžaloby

Případ R.Cannan

„Šance, že biologická stopa patří někomu jinému, je 1 : 260 mil.“

Logický trik obhajoby

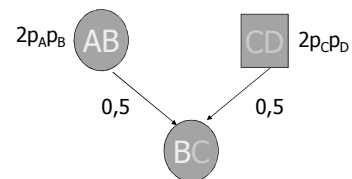
Výpočet čitatele na příkladu



Výpočet čitatele X - pravděpodobnost, že

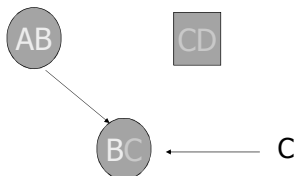
- žena náhodně vybraná z populace je AB,
- muž náhodně vybraný z populace je CD,
- jejich dítě je BC.

Výpočet čitatele na příkladu



$$X = 2p_{AP_B} * 2p_{CP_D} * 0,5 * 0,5$$

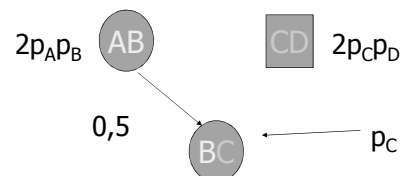
Výpočet jmenovatele na příkladu



Výpočet jmenovatele Y - pravděpodobnost, že

- žena náhodně vybraná z populace je AB,
- muž náhodně vybraný z populace je CD,
- dítě této ženy a jiného neznámého muže je BC.

Výpočet jmenovatele na příkladu



$$\text{Pravděpodobnost} = 2p_{AP_B} * 2p_{CP_D} * 0,5 * p_C$$

Index otcovství

$$PI = \frac{\cancel{2p_A p_B} * \cancel{2p_C p_D} * 0,5 * 0,5}{\cancel{2p_A p_B} * \cancel{2p_C p_D} * 0,5 * p_C} =$$
$$= \frac{0,5}{p_C}$$

Nesprávný popis zlomku

- Čítec je pravděpodobnost, že testovaný muž je otcem dítěte.
- Jmenovatel je pravděpodobnost, že testovaný muž není otcem dítěte.

Správný popis zlomku

- Čítec je pravděpodobnost výsledků genotypizace (pravděpodobnost rodokmene) za předpokladu, že testovaný muž je biologickým otcem dítěte.
- Jmenovatel je pravděpodobnost výsledků genotypizace (pravděpodobnost rodokmene) za předpokladu, že biologickým otcem je neznámý náhodný muž z populace.

Nesprávné slovní vyjádření indexu otcovství

- Je (X/Y) krát pravděpodobnější, že testovaný muž je biologickým otcem dítěte, než že neznámý náhodný muž je otcem dítěte.

Správné slovní vyjádření indexu otcovství (1)

Genotypizační výsledky jsou (PI=X/Y) krát pravděpodobnější za předpokladu, že testovaný muž je otcem dítěte, než pokud neznámý náhodný muž z populace je otcem dítěte.

Správné slovní vyjádření indexu otcovství (2)

Podpora genetických výsledků je (PI=X/Y) krát větší, pokud testovaný muž je biologickým otcem, než pokud náhodný neznámý muž z populace je biologickým otcem.

Fráze pro vyjádření LR

„Ať jsou jiné důkazy ve prospěch H_1 jakkoli velké, toto vyšetření je posiluje LR-krát.“

„Selský rozum“ a index otcovství

Fakt, že testovaný muž má alelu biologického otce může být důkazem *proti* otcovství testovaného (důkazem snižujícím PO).

„Selský rozum“ a index otcovství

M	D	Ž
pq	rq	rp

$r = 0,6; \quad p = 0,2; \quad q = 0,2$

$$LR = \frac{1}{2 \times r} = \frac{1}{2 \times 0,6} = 0,8$$

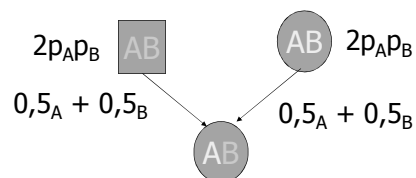
Možné genotypy tria M-D-Ž

- 15 možných rodokmenů
- Rodokmen 1 (M-D-N: AB-BC-CD) jsme již probrali na příkladu

Rodokmen 2

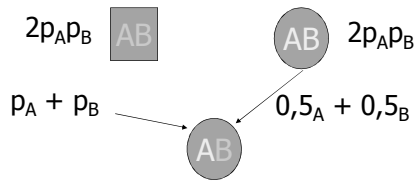
- Matka, dítě a žalovaný muž mají všichni stejnou kombinaci alel (AB)
- Matka má 50% šanci předat alelu A a 50% šanci předat alelu B
- Nařčený má 50% šanci předat alelu A a 50% šanci předat alelu B
- Náhodný muž z populace má $(p_A + p_B) * 100\%$ šanci předat alelu A nebo B
- Odvodíme si, že $PI = 1/(p_A + p_B)$

Čítatel X pro rodokmen 2



$$X = 2p_A p_B * 2p_A p_B * (0,5_A * 0,5_B + 0,5_B * 0,5_A)$$

Jmenovatel Y pro rodokmen 2



$$Y = 2p_{AP_B} * 2p_{AP_B} * (0,5_A * p_B + 0,5_B * p_A)$$

Index otcovství (rodokmen 2)

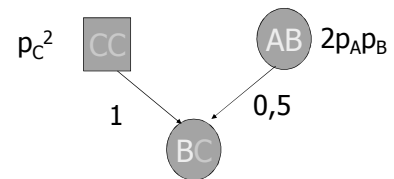
$$PI = \frac{2p_{AP_B} * 2p_{AP_B} * (0,5_A * 0,5_B + 0,5_B * 0,5_A)}{2p_{AP_B} * 2p_{AP_B} * (0,5_A * p_B + 0,5_B * p_A)} =$$

$$= \frac{0,5}{0,5_A * p_B + 0,5_B * p_A} = \frac{1}{p_B + p_A}$$

Rodokmen 3

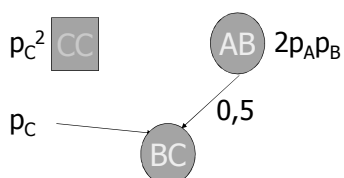
- Matka AB, Dítě BC, Nařčený CC
- Nařčený může předat jen alelu C
- Náhodný muž má p_C šanci předat alelu
- Odvoďte PI

Číselník X pro rodokmen 3



$$X = 2p_{AP_B} * p_C^2 * 0,5 * 1$$

Jmenovatel Y pro rodokmen 3



$$Y = 2p_{AP_B} * p_C^2 * 0,5 * p_C$$

Index otcovství (rodokmen 3)

$$PI = \frac{2p_{AP_B} * p_C^2 * 0,5 * 1}{2p_{AP_B} * p_C^2 * 0,5 * p_C} =$$

$$= \frac{1}{p_C}$$

Rodokmen 4

- Matka AB, Dítě AB, Nařčený BC
- Matka má 50% šanci předat A a 50% šanci předat alelu B
- Nařčený má 50% šanci předat alelu B
- Náhodný muž má $p_A + p_B$ šanci předat alelu A nebo B
- Odvoďte PI

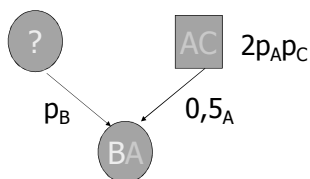
PI

$$X = P(\text{žalovaný předá A}) \cdot p_B + P(\text{žalovaný předá B}) \cdot p_A = 0,5 \cdot p_B + 0 \cdot p_A = 0,5 \cdot p_B$$

$$Y = p_A \cdot p_B + p_B \cdot p_A = 2p_A p_B$$

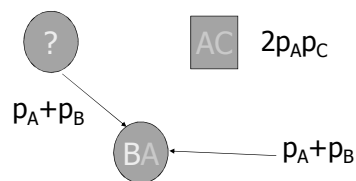
$$PI = \frac{0,5 \cdot p_B}{2p_A p_B} = \frac{1}{4p_A}$$

Rodokmen bez matky 1



$$X = 2p_A p_C \cdot 0,5 \cdot p_B$$

Výpočet jmenovatele rodokmene bez matky 1



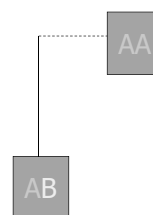
$$Y = 2p_A p_C \cdot (p_A \cdot p_B + p_B \cdot p_A)$$

PI

$$PI = \frac{\cancel{2p_A p_C} \cdot 0,5 \cdot p_B}{\cancel{2p_A p_C} \cdot (p_A \cdot p_B + p_B \cdot p_A)} =$$

$$= \frac{0,5 \cdot \cancel{p_B}}{\cancel{2p_A} \cdot \cancel{p_B}} = \frac{1}{4p_A}$$

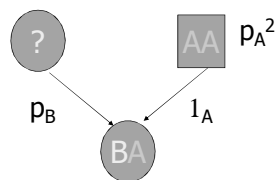
Rodokmen bez matky 2



Rodokmen bez matky 2

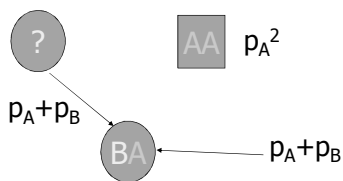
- Matka mohla předat A nebo B
- Žalovaný mohl předat jen A
- Náhodný muž má $p_A + p_B$ pravděpodobnost předání alely A nebo B

Výpočet čitatele rodokmene bez matky 2



$$X = p_A^2 * 1_A * p_B$$

Výpočet jmenovatele rodokmene bez matky 2

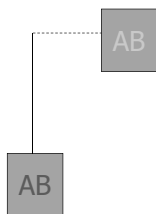


$$Y = p_A^2 * (p_A * p_B + p_B * p_A)$$

PI

$$PI = \frac{p_A^2 * 1_A * p_B}{p_A^2 * (p_A * p_B + p_B * p_A)} = \frac{p_B}{2p_A * p_B} = \frac{1}{2p_A}$$

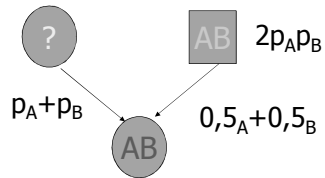
Rodokmen bez matky 3



Rodokmen bez matky 3

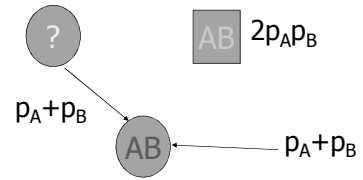
- Matka mohla předat A nebo B
- Žalovaný mohl předat A nebo B
- Náhodný muž má $p_A + p_B$ pravděpodobnost předání alely A nebo B

Výpočet čitatele rodokmene bez matky 3



$$X = 2p_{APB} * (0,5 p_B + 0,5 p_A)$$

Výpočet jmenovatele rodokmene bez matky 3



$$Y = 2p_{APB} * (p_A * p_B + p_B * p_A)$$

PI

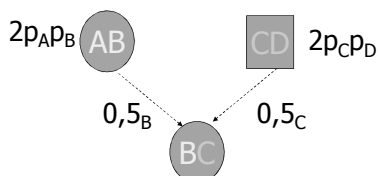
$$PI = \frac{2p_{APB} * (0,5 p_B + 0,5 p_A)}{2p_{APB} * (p_A * p_B + p_B * p_A)} =$$

$$= \frac{0,5 p_B + 0,5 p_A}{2p_{APB}} = \frac{p_A + p_B}{4p_{APB}}$$

Reverzní paternita

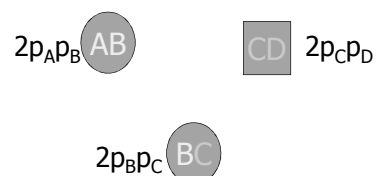
- Genotypy
 - matka
 - nalezené dítě
 - otec
- X = pravděpodobnost, že náhodná matka má genotyp AB a náhodný muž má genotyp CD a jejich dítě má genotyp BC
- Y = pravděpodobnost, že náhodná žena nepřibuzná dítěti má genotyp AB a náhodný muž nepřibuzný dítěti má genotyp CD a náhodné dítě má genotyp BC

Výpočet čitatele při reverzní paternitě 1



$$X = 2p_{APB} * 2p_{CPD} * 0,5 p_B * 0,5 p_C$$

Výpočet jmenovatele při reverzní paternitě 1



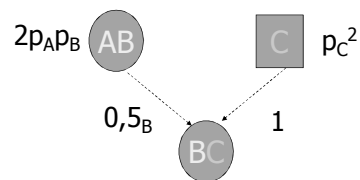
$$Y = 2p_{APB} * 2p_{CPD} * 2p_{BPC}$$

PI

$$LR = \frac{\cancel{2p_{AP_B}} * \cancel{2p_{CP_D}} * 0,5_B * 0,5_C}{\cancel{2p_{AP_B}} * \cancel{2p_{CP_D}} * 2p_{BP_C}} =$$

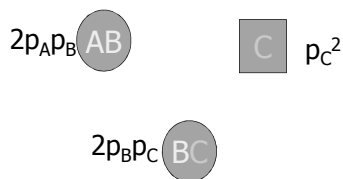
$$= \frac{1}{8p_{BP_C}}$$

Výpočet čitatele při reverzní paternitě 2



$$X = 2p_{AP_B} * p_C^2 * 0,5_B * 1_C$$

Výpočet jmenovatele při reverzní paternitě 2



$$Y = 2p_{AP_B} * p_C^2 * 2p_{BP_C}$$

PI

$$LR = \frac{\cancel{2p_{AP_B}} * \cancel{p_C^2} * 0,5_B * 1_C}{\cancel{2p_{AP_B}} * \cancel{p_C^2} * 2p_{BP_C}} =$$

$$= \frac{1}{4p_{BP_C}}$$

- V reálných případech reverzní paternity nebývají k typizaci dostupní oba rodiče.
- Pokud je jen jeden rodič ke genotypizaci, tak je případ obdobou výpočtu LR „bez matky“.

Děkuji vám za pozornost!