

Vaterschaftstatistik in Erbsystemen mit N kodominanten Allelen

Jürgen Conradt

Dr. Jürgen Conradt, Zasius-Straße 91, D-7800 Freiburg

BEZEICHNUNGEN

Wir betrachten ein System mit N kodominanten Allelen. Die Allele nummerieren wir von 1 bis N durch und sprechen sie durch ihre Nummer an. Als Variable für diese Allelnummer verwenden wir i, j, v, m . Einen Genotyp, bestehend aus den Allelen i und j , notieren wir durch $i-j$. Für einen Genotyp $i-j$ definieren wir eine Funktion d_{i-j} , welche 1 ist für i gleich j (homozygoter Genotyp) und 2 für i ungleich j (heterozygoter Genotyp).

AUSWERTUNG DES KIND-MUTTER-VATER-FALLES

Bei der Auswertung eines KMV-Falles unterscheiden wir die folgenden Kombinationen:

T1 das Kind ist homozygot;

T2 das Kind ist heterozygot und

T2.1 Kind, Mutter und Vater tragen den gleichen heterozygoten Genotyp,

T2.2 nur Kind und Mutter tragen den gleichen heterozygoten Genotyp, einen davon verschiedenen trägt der Vater,

T2.3 Kind und Mutter tragen verschiedene Genotypen, der Genotyp des Vaters ist beliebig.

Für diese Möglichkeiten berechnen wir in Tabelle 1 die Wahrscheinlichkeiten für die Kind-Mutter-Dublette $p(KM)$, für den Phänotyp des Vaters $p(V)$, für $Y=p(KM)p(V)$ und $X=p(KMV)$. Die Probanden Mutter M, unbekannter Erzeuger U und der Vater V entstammen möglicherweise verschiedenen Populationen. Die Wahrscheinlichkeit eines Allels i bezeichnen wir deshalb mit M_i , U_i oder V_i . Die letzte Spalte der Tabelle 1 enthält den Likelihoodquotienten $L=Y/X$. Wie wir sehen, spielt die Population der Mutter M und des Vaters V keine Rolle. Der Likelihoodquotient L wird ausschließlich gebildet mit den Frequenzen des unbekanntes Erzeugers U. Diese alleinige Abhängigkeit von der Population des unbekanntes Erzeugers ist weithin nicht bekannt. Mit den Formeln der letzten Spalte von Tabelle 1 ist es leicht, Befunde einer KMV-Terzette auszuwerten.

Tabelle 1. X, Y und L für den KMV-Fall

	M	V	Y=p(KM)p(V)		X	L
	K		p(KM)	p(V)	p(KMV)	Y/X
T1	i-m i-i	i-v	M _i M _m U _i	d _{i-v} V _i V _v	M _i M _m V _i V _v	d _{i-v} U _i
T2.1	i-j i-j i≠j	i-j	M _i M _j U _i + M _i M _j U _j	2V _i V _j	2M _i M _j V _i V _j	U _i +U _j
T2.2	i-j i-j i≠j, i≠v	j-v	M _i M _j U _i + M _i M _j U _j	d _{j-v} V _j V _v	M _i M _j V _j V _v	d _{j-v} (U _i +U _j)
T2.3	i-m i-j i≠j, j≠m	j-v	M _i M _m U _j	d _{j-v} V _j V _v	M _i M _m V _j V _v	d _{j-v} U _j

LEISTUNGSKENNZIFFERN FÜR DEN KIND-MUTTER-VATER-FALL

Einführung

Die Erwartungswerte biostatistischer Kennziffern eines Falltyps kennzeichnen die Leistung eines Blutgruppensystems. Wir betrachten nur den KMV-Fall und exemplarisch als Kennziffer den Likelihoodquotienten L. Für die Berechnung des Erwartungswertes des Likelihoodquotienten L müssen wir jeden Wert L(KMV) mit der Wahrscheinlichkeit seines Auftretens gewichten. Die Wahrscheinlichkeit einer KMV-Terzette ist unter den Hypothese H_X bzw. H_Y verschieden. Wir notieren die bedingten Erwartungswerte mit E_{KMV}(L|H_X) resp. E_{KMV}(L|H_Y). Diese Erwartungswerte werden durch Summation Σ* über alle genetisch möglichen KMV-Terzette berechnet.

$$E_{KMV}(L|H_X) = \Sigma^* X(KMV) L(KMV) = \Sigma^* XY/X = \Sigma^* Y = \bar{A} = E_{KMV}(\bar{A})$$

$$E_{KMV}(L|H_Y) = (\Sigma^* Y(KMV) L(KMV)) / \Sigma^* Y(KMV) = (\Sigma^* Y^2/X) / \bar{A}$$

Die Summe Σ*X(KMV) ergibt Eins. Die Summe Σ*Y(KMV) = Σ*p(KM)p(V) ist die mittlere Nichtausschlußwahrscheinlichkeit \bar{A} . Da diese i.a. ungleich Eins ist, müssen wir die Gewichte Y(KMV) durch Division mit \bar{A} normieren.

Die obigen Aussagen gelten für beliebige Erbsysteme.

Numerische Durchführung

Zur numerischen Berechnung der Erwartungswerte können wir alle KMV-Terzette aufstellen, auf genetische Kompatibilität prüfen und auswerten. Aber wie wir Tabelle 1 entnehmen, gibt es in kodominanten Systemen unter der Annahme der gleichen Population P_i für alle Probanden nur vier verschieden gebildete Likelihoodquotienten (Tabelle 2).

Tabelle 2. Die verschiedenen Likelihoodquotienten

Typ	Likelihood	Kombination von Tabelle 1
I	$L = P_i$	T1 und T2.3
II	$L = 2P_i$	T1 und T2.3
III	$L = P_i + P_j$	T2.1
IV	$L = 2(P_i + P_j)$	T2.2

In einer längeren Handrechnung fassen wir die Gewichte für X bzw. Y dieser vier Likelihoodquotienten zusammen und erhalten Tabelle 3.

Tabelle 3. Formeln zur Berechnung der Leistungskennziffern

Typ	Summe	Gewichte für Y	Gewichte für X	L
I	$\sum_{i=1}^N$	$P_i^3(1-P_i+P_i^2)$	$P_i^2(1-P_i+P_i^2)$	P_i
II	$\sum_{i=1}^N$	$2P_i^2(1-P_i)(1-P_i+P_i^2)$	$P_i(1-P_i)(1-P_i+P_i^2)$	$2P_i$
III	$\sum_{i=1}^{N-1} \sum_{j=i+1}^N$	$P_i P_j (P_i + P_j)^3$	$P_i P_j (P_i + P_j)^2$	$P_i + P_j$
IV	$\sum_{i=1}^{N-1} \sum_{j=i+1}^N$	$2P_i P_j (P_i + P_j)^2 (1 - P_i - P_j)$	$P_i P_j (P_i + P_j) (1 - P_i - P_j)$	$2(P_i + P_j)$
	$N(N+1)$	$\Sigma^*Y = \bar{A}$	$\Sigma^*X = 1$	

Die Formeln aus der Tabelle 3 ermöglichen eine schnelle Berechnung beliebiger Leistungskennziffern. Dabei müssen nur $N(N+1)$ Produkte berechnet und summiert werden.

OPTIMALE GESTALT EINES SYSTEMS

Optimale Verteilung der Allelwahrscheinlichkeiten

Wir beschränken uns im folgenden auf Leistungskennziffern für den KMV-Fall. Wie schon bekannt, z.B. Selvin (1980), ist die Ausschlußwahrscheinlichkeit in einem kodominanten Erbsystem am größten, wenn die Allele gleichwahrscheinlich sind. Entsprechend haben $E_{KMV}(L|H_X) = \bar{A}$ und $E_{KMV}(L|H_Y)$ bei gleichverteilten Allelen ein Minimum.

Optimale Anzahl der Allele

Wir stellen die Frage, ob es im biostatistischen Sinne günstiger ist, ein System mit 'vielen' Allelen oder mehrere Systeme mit 'wenigen' Allelen zu verwenden. Dazu setzen wir eine Gleichverteilung der Allelwahrscheinlichkeiten voraus, $P_i = 1/N$, und berechnen den Beitrag eines Alleles zur Leistungskennziffer mit

$$L_X = \sqrt[N]{E_{KMV}(L|H_X)}, \quad L_Y = \sqrt[N]{E_{KMV}(L|H_Y)} .$$

Für verschiedene Werte von N erhalten wir Tabelle 4.

Tabelle 4. Beitrag eines Alleles zur biostatistischen Leistung

N	2	3	4	5	6	7	8
L_X	.9014	.8571	.8392	<u>.8345</u>	.8355	.8391	.8436
L_Y	.9405	.8986	.8685	.8559	<u>.8518</u>	.8519	.8541

Die optimale Anzahl Allele in einem System mit Gleichverteilung der Allelwahrscheinlichkeiten ist 5 für $E_{KMV}(L|H_X)$ und 6 für $E_{KMV}(L|H_Y)$. Angenommen wir hätten die Wahl zwischen einem System mit N=15 Allelen und drei Systemen mit je N=5 Allelen, so erhalten wir die Resultate der Tabelle 5.

Tabelle 5. Vergleich der Kennziffern

	N=5	(N=5) ³	N=15
$E_{KMV}(L H_X) = \bar{A}$	0.4048	0.0663	0.1364
$E_{KMV}(L H_Y)$	0.4593	0.0969	0.1454

Der Erwartungswert $E_{KMV}(L|H_X) = \bar{A}$ den Putativ-Vater nicht auszuschließen, ist mit den drei Systemen ($\bar{A}=6.6\%$) nur halb so groß wie in dem einem System mit 15 Allelen ($\bar{A}=13.6\%$).

LITERATUR

- Ohno Y, Sebetan IM, Akaishi S (1982) A simple method for calculating the probability of excluding paternity with any number of codominant alleles. Forensic Science International 19: 93-98
- Selvin S (1980) Probability of nonpaternity determined by multiple allele codominant systems. Am J Hum Genet 32:276-278
- Szczotka H, Hummel K (1976) Calculation of the plausibility of paternity with blood group results using gene frequencies of foreign populations: Part I: Formulas for the manual computation of Y/X values. Z Immun Forsch 151:359-373
- Wiener AS (1968) Chances of proving nonpaternity with a system determined by triple allelic codominant genes. Am J Hum Genet 20:279-282