Affected Sib-Pair Two-Locus Linkage Analysis

Algebraic Statistical Genetics

Ingileif Bryndís Hallgrímsdóttir

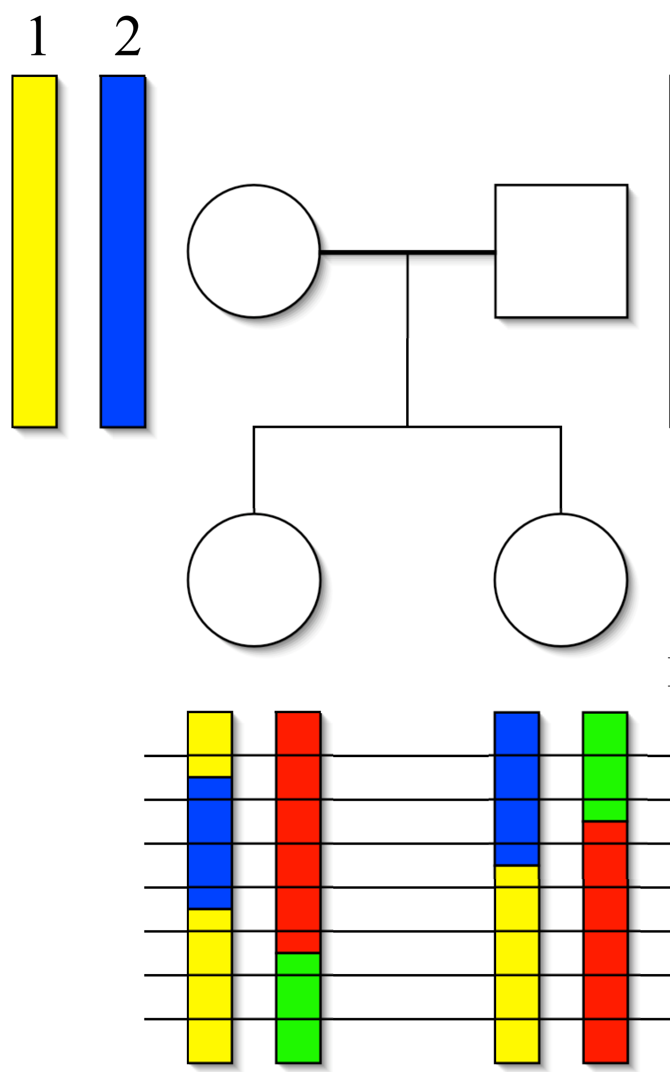
Department of Statistics

University of California, Berkeley

Outline

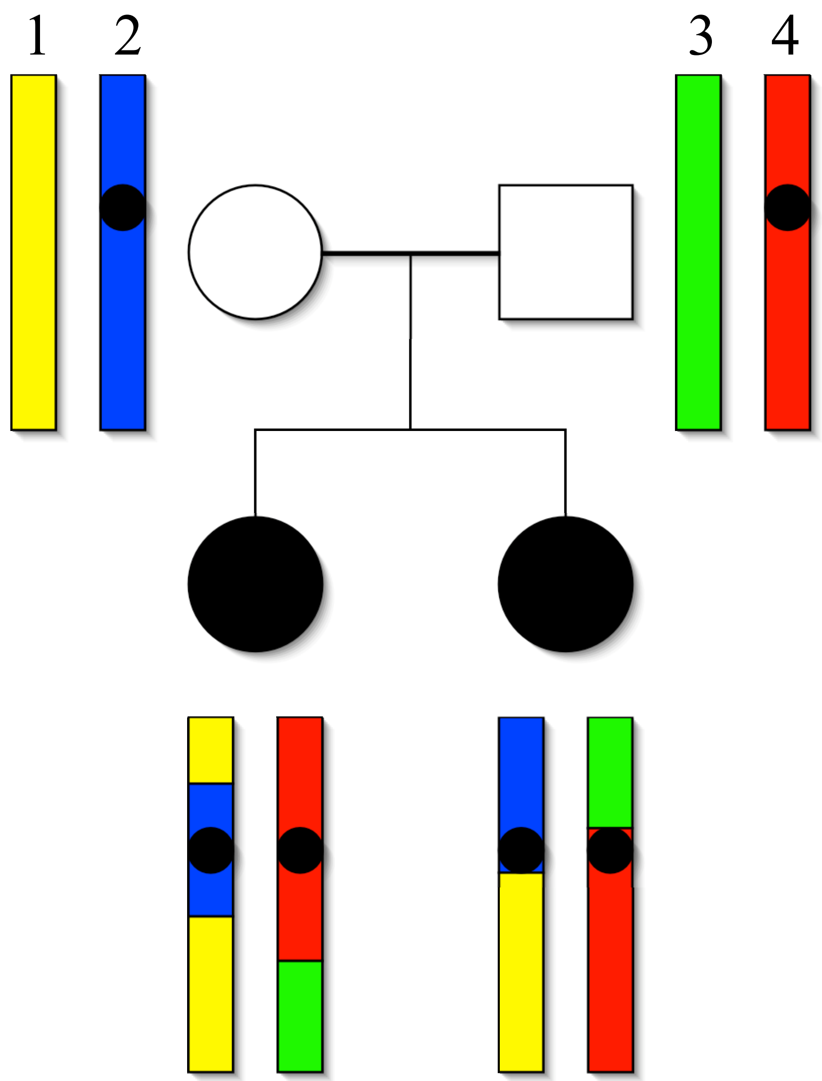
- ◇ Brief introduction to Linkage Analysis.
- ◇ Review of the one locus model, some geometry and algebra.
- ◇ Components of the log-likelihood ratio.
- ◇ The two locus model.
- ◇ Four new correlation statistics to test for interactions.

Identity by Descent Sharing

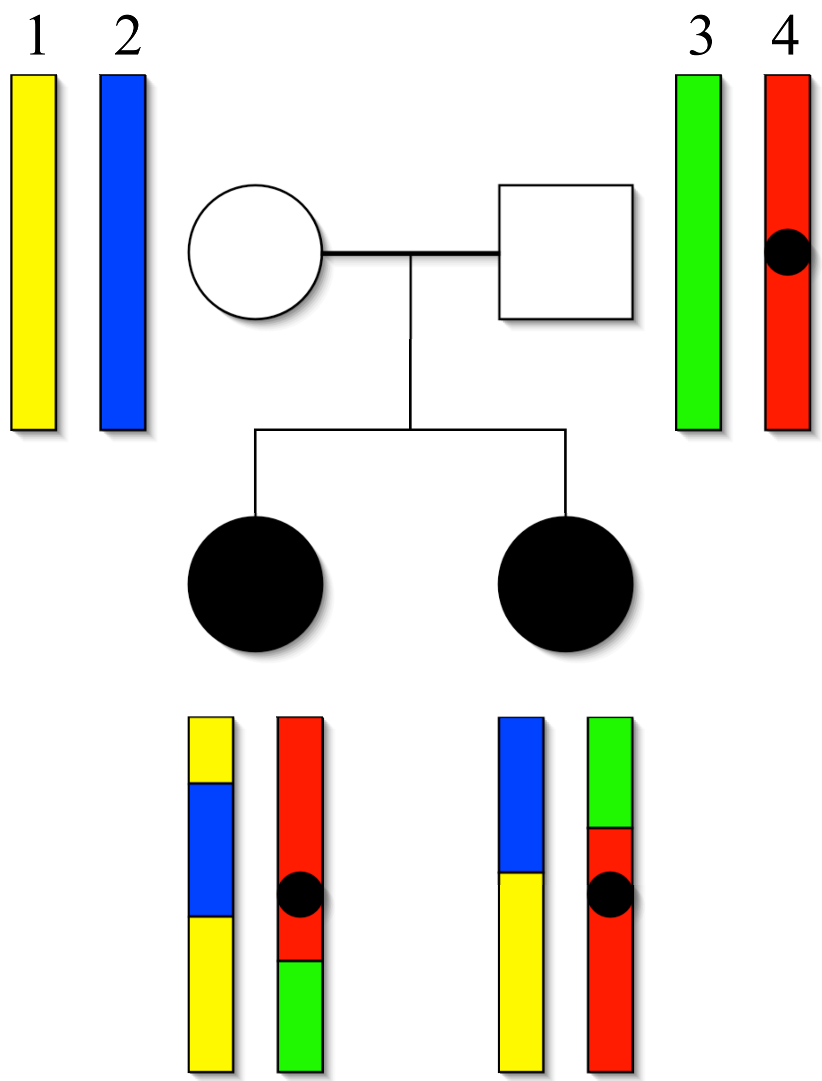


IBD	Inheritance vector
0	(1,4,2,3)
1	(2,4,2,3)
2	(2,4,2,4)
1	(2,4,1,4)
2	(1,4,1,4)
1	(1,3,1,4)
1	(1,3,1,4)

Increased IBD Sharing - Recessive Model



Increased IBD Sharing - Dominant Model



The Data

We collect affected sib-pairs and genotype them at markers across the genome.

	n_0	n_1	n_2
Marker 1	26	49	25
Marker 2	20	48	32
Marker 3	8	50	42
⋮			
Marker m	27	53	20

n_i: #affected sib-pairs that share i chromosomes IBD, $i = 0, 1, 2$

Disease Models

$$f_{dd} = Pr(\text{Affected}|\text{Genotype is } dd),$$
$$f_{Dd} = Pr(\text{Affected}|\text{Genotype is } Dd),$$
$$f_{DD} = Pr(\text{Affected}|\text{Genotype is } DD).$$

		f_{dd}	f_{Dd}	f_{DD}
strict-recessive	:	0	0	f
quasi-recessive	:	rf	rf	f
strict-dominant	:	0	f	f
quasi-dominant	:	rf	f	f
additive	:	rf	$\frac{r+1}{2}f$	f

where $f, r \in [0, 1]$.

The disease allele is D and its population frequency is p.

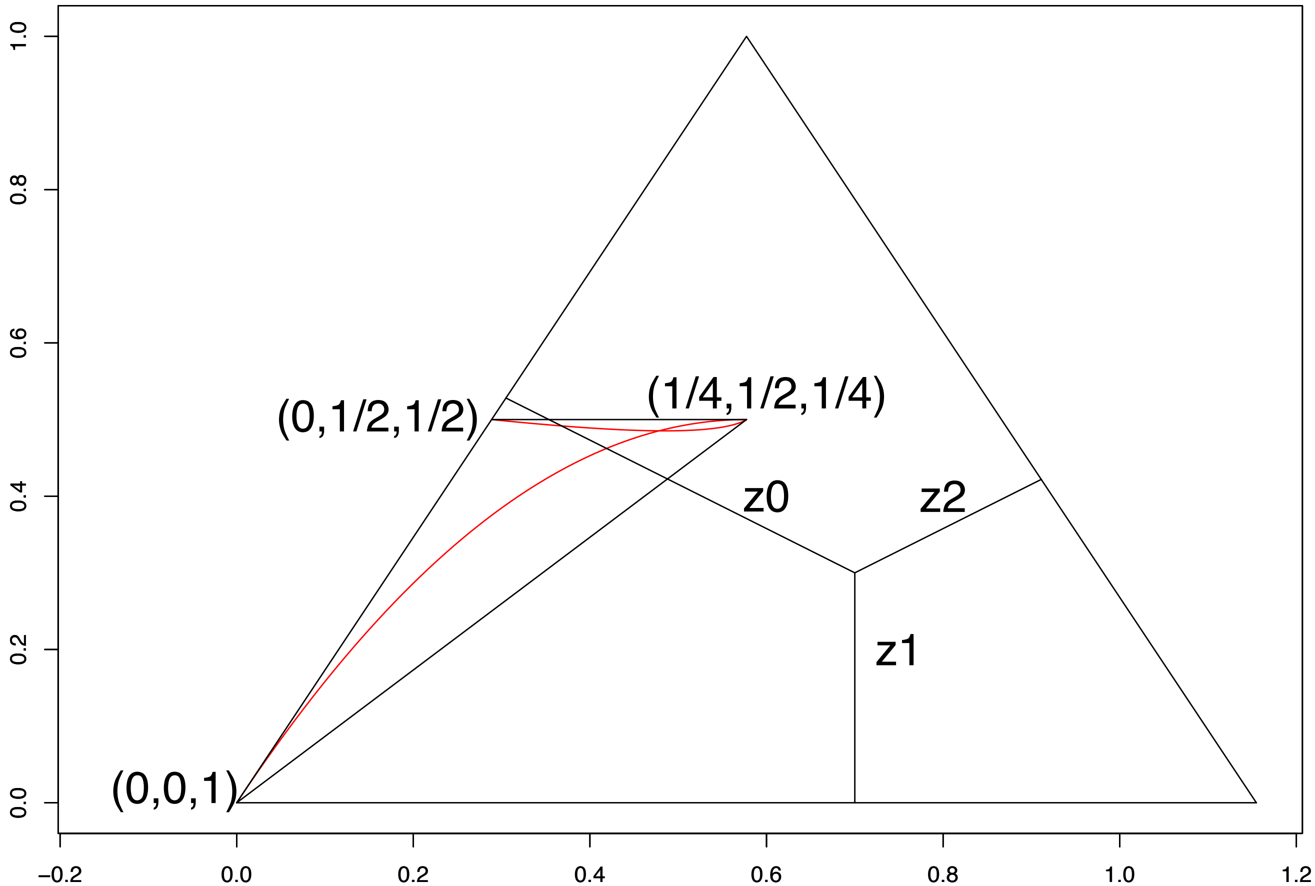
Affected Sib-Pair (ASP) IBD Probabilities

It is easy to show that

$$\begin{aligned} z_i &= Pr(IBD = i|ASP) \quad i = 0, 1, 2 \\ &\propto \sum_{\substack{\text{inheritance} \\ \text{vectors}}} \sum_{\substack{\text{parental} \\ \text{genotypes}}} (1-p)^j p^{4-j} f_{a_1 a_2} f_{b_1 b_2} \end{aligned}$$

Where a_1a_2 and b_1b_2 are the the genotypes of the first and second sibling respectively and j is the total number of disease genes among the parents.

Holmans' trinagle I



Defining Polynomials

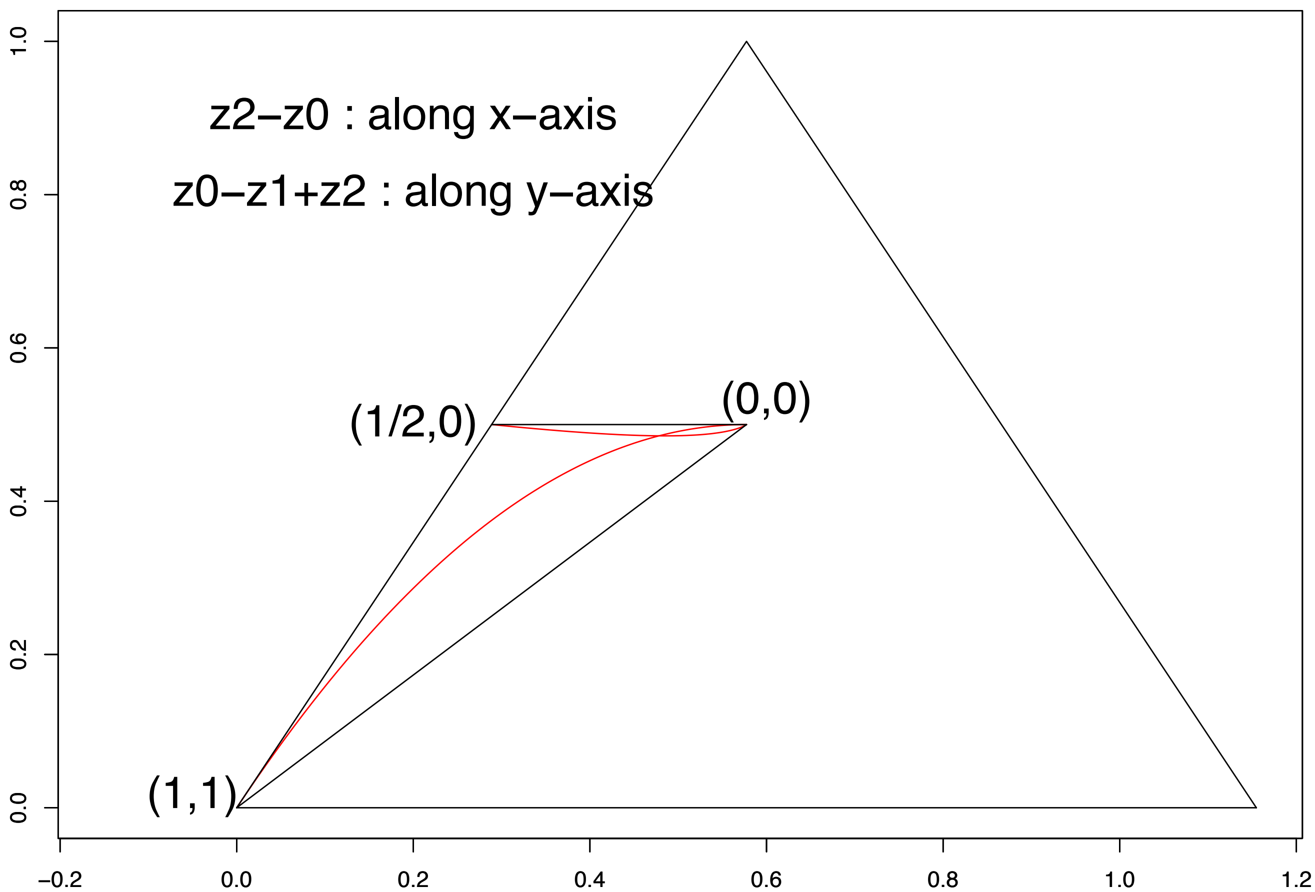
Recessive (Hardy-Weinberg)

$$\begin{aligned} z_1^2 - 4z_0 z_2 &= 0 \\ &\Updownarrow \\ (z_0 - z_1 + z_2) - (z_2 - z_0)^2 &= 0 \end{aligned}$$

Additive

$$\begin{aligned} z_1 &= \frac{1}{2} \\ &\Updownarrow \\ z_0 - z_1 + z_2 &= 0 \end{aligned}$$

Holmans' triangle II



Defining Polynomial

Dominant

$$z_0 z_1^2 + 4z_0^2 z_2 - 8z_0 z_1 z_2 + 4z_0 z_2^2 + 4z_1 z_2^2 - 4z_2^3 = 0$$

Defining Polynomial

```
(-64*f0*f1^5*f2^2-32*f0^2*f1^6+64*f0*f1^6*f2-16*f1^4*f2^4+64*f0^3*f1^5-64*f0^2*f1*f2^5
-64*f0^5*f1*f2^2-32*f0^5*f2^3-16*f0^4*f1^4+64*f0^3*f1*f2^4+32*f0^4*f2^4-32*f0^3*f2^5
+64*f0^4*f1*f2^3+16*f0^2*f2^6-32*f1^6*f2^2+64*f1^5*f2^3+16*f0^6*f2^2+160*f0^4*f1^2*f2^2
+160*f0^2*f1^2*f2^4-96*f0^3*f1^4*f2+224*f0^2*f1^4*f2^2-64*f0^2*f1^5*f2-96*f0*f1^4*f2^3
-320*f0^3*f1^2*f2^3) * z0^3 * z2  +

(-16*f1^5*f2^3+4*f1^4*f2^4+32*f0^3*f1^4*f2+32*f0*f1^7-16*f1^8+32*f1^7*f2+16*f0^5*f1*f2^2
+4*f0^4*f1^4-4*f0^2*f2^6+16*f0^3*f1*f2^4-8*f0^4*f2^4+32*f0*f1^4*f2^3-32*f0^2*f1^3*f2^3
-16*f0^3*f1^5+16*f0^2*f1*f2^5-32*f0^3*f1^3*f2^2-16*f0^2*f1^5*f2+88*f0^2*f1^4*f2^2
-32*f0^2*f1^2*f2^4-16*f0*f1^5*f2^2-32*f0^4*f1^2*f2^2-4*f0^6*f2^2+16*f0^4*f1*f2^3
-64*f0*f1^6*f2) *z0^2 * z1^2  +

(-192*f0*f1^6*f2-128*f0^3*f1^3*f2^2+32*f0^2*f1*f2^5-256*f0^3*f1*f2^4-128*f0^2*f1^3*f2^3
+32*f0*f1^3*f2^4+32*f1^8+16*f0^4*f1^4+32*f0^5*f1*f2^2+32*f0^3*f1^4*f2-128*f0^4*f1^2*f2^2
+96*f0^5*f2^3-64*f0^4*f2^4-64*f0^3*f1^5+96*f0^3*f2^5-16*f0^2*f2^6-32*f0*f1^7+32*f0^4*f1^3*f2
+960*f0^3*f1^2*f2^3-128*f0^2*f1^2*f2^4-736*f0^2*f1^4*f2^2+416*f0^2*f1^5*f2-256*f0^4*f1*f2^3
+416*f0*f1^5*f2^2+32*f0*f1^4*f2^3-64*f1^5*f2^3+16*f1^4*f2^4-32*f1^7*f2-16*f0^6*f2^2)
* z0^2 * z1 *z2  +

(-128*f0^2*f1^4*f2^2+64*f0^2*f1*f2^5-160*f0^3*f1^4*f2-16*f1^8+768*f0^3*f1^3*f2^2
-192*f0^2*f1^5*f2+64*f0^5*f1*f2^2+768*f0^2*f1^3*f2^3-16*f0^4*f1^4+48*f0^4*f2^4
+320*f0^3*f1*f2^4-96*f0^5*f2^3+64*f1^5*f2^3-1216*f0^3*f1^2*f2^3-192*f0^4*f1^2*f2^2
+320*f0^4*f1*f2^3-16*f1^4*f2^4-96*f0^3*f2^5-192*f0*f1^5*f2^2+64*f0^3*f1^5-160*f0*f1^4*f2^3
+192*f0*f1^6*f2-192*f0^2*f1^2*f2^4) * z0^2 * z2^2  +

(128*f0*f1^6*f2+72*f0^3*f1^3*f2^2-8*f0^2*f1*f2^5+8*f0^3*f1*f2^4+72*f0^2*f1^3*f2^3
-8*f0*f1^3*f2^4+32*f1^8-4*f0^4*f1^4-8*f0^5*f1*f2^2-24*f0^3*f1^4*f2+16*f0^4*f1^2*f2^2
-8*f0^5*f2^3+8*f0^4*f2^4+16*f0^3*f1^5-8*f0^3*f2^5+4*f0^2*f2^6-64*f0*f1^7-8*f0^4*f1^3*f2
-64*f0^3*f1^2*f2^3+16*f0^2*f1^2*f2^4-104*f0^2*f1^4*f2^2-16*f0^2*f1^5*f2+8*f0^4*f1*f2^3
-16*f0*f1^5*f2^2-24*f0*f1^4*f2^3+16*f1^5*f2^3-4*f1^4*f2^4+16*f1^6*f2^2-64*f1^7*f2+4*f0^6*f2^2
```

```
+16*f0^2*f1^6) * z0 * z1^3  +

(64*f0*f1^6*f2-160*f0^3*f1^3*f2^2-16*f0^2*f1*f2^5+48*f0^3*f1*f2^4-160*f0^2*f1^3*f2^3
-32*f0*f1^3*f2^4-96*f1^8+4*f0^4*f1^4-16*f0^5*f1*f2^2+120*f0^3*f1^4*f2+128*f0^4*f1^2*f2^2
-40*f0^5*f2^3+72*f0^4*f2^4-16*f0^3*f1^5-40*f0^3*f2^5+4*f0^2*f2^6+128*f0*f1^7-32*f0^4*f1^3*f2
-352*f0^3*f1^2*f2^3+128*f0^2*f1^2*f2^4+872*f0^2*f1^4*f2^2-464*f0^2*f1^5*f2+48*f0^4*f1*f2^3
-464*f0*f1^5*f2^2+120*f0*f1^4*f2^3-16*f1^5*f2^3+4*f1^4*f2^4+16*f1^6*f2^2+128*f1^7*f2
+4*f0^6*f2^2+16*f0^2*f1^6) * z0 * z1^2 * z2  +

(768*f0^3*f1^2*f2^3-32*f0^4*f1*f2^3-32*f0^5*f1*f2^2+64*f0^5*f2^3+96*f1^8+32*f0^4*f1^3*f2
-160*f0^4*f2^4-32*f0^3*f1*f2^4+576*f0^2*f1^5*f2-320*f0^2*f1^4*f2^2-32*f0^2*f1*f2^5
-480*f0^2*f1^3*f2^3+576*f0*f1^5*f2^2+32*f0*f1^3*f2^4-384*f0*f1^6*f2-64*f0*f1^7-64*f1^6*f2^2
-64*f1^7*f2+64*f0^3*f2^5-480*f0^3*f1^3*f2^2-64*f0^2*f1^6) * z0 * z1 * z2^2  +

(32*f1^6*f2^2-64*f0^3*f1^2*f2^3-32*f1^8+96*f0^4*f2^4-64*f0*f1^4*f2^3+192*f0*f1^6*f2
+32*f0^2*f1^2*f2^4+32*f0^4*f1^2*f2^2-128*f0^4*f1*f2^3-64*f0^3*f1^4*f2+256*f0^3*f1^3*f2^2
-128*f0*f1^5*f2^2+32*f0^2*f1^6-128*f0^3*f1*f2^4-128*f0^2*f1^5*f2+256*f0^2*f1^3*f2^3
-192*f0^2*f1^4*f2^2) * z0 * z2^3  +

(32*f1^7*f2-16*f1^8-8*f0^4*f1^2*f2^2-8*f0^3*f1^3*f2^2+4*f0^5*f2^3+32*f0*f1^7+32*f0*f1^5*f2^2
-64*f0*f1^6*f2-8*f0^2*f1^2*f2^4+8*f0^4*f1^3*f2+32*f0^2*f1^5*f2+16*f0^3*f1^2*f2^3-16*f1^6*f2^2
+8*f0*f1^3*f2^4-16*f0^2*f1^6-f0^6*f2^2-f0^2*f2^6-6*f0^4*f2^4-8*f0^3*f1^4*f2-8*f0*f1^4*f2^3
+4*f0^3*f2^5-8*f0^2*f1^3*f2^3) * z1^4  +

(64*f1^8-16*f0*f1^4*f2^3-8*f0^3*f1*f2^4+64*f0*f1^5*f2^2-16*f0^2*f1^2*f2^4+64*f0^2*f1^5*f2
+8*f0^2*f1*f2^5+32*f0^3*f1^2*f2^3+40*f0^2*f1^3*f2^3+32*f1^6*f2^2+40*f0^3*f1^3*f2^2
-16*f0^4*f1^2*f2^2-8*f0*f1^3*f2^4-96*f1^7*f2-16*f0^3*f1^4*f2-8*f0^4*f1^3*f2+8*f0^5*f1*f2^2
+32*f0^2*f1^6+64*f0*f1^6*f2-160*f0^2*f1^4*f2^2-8*f0^4*f1*f2^3-96*f0*f1^7) * z1^3 * z2  +

(24*f0^3*f1^4*f2+96*f0^2*f1^3*f2^3+24*f0*f1^4*f2^3-192*f0^2*f1^5*f2+96*f0^3*f1^3*f2^2
-16*f1^6*f2^2-8*f0^3*f2^5+96*f0*f1^7-96*f1^8+96*f1^7*f2-16*f0^2*f1^6+16*f0^4*f2^4-8*f0^5*f2^3
-192*f0*f1^5*f2^2-96*f0^3*f1^2*f2^3+48*f0^2*f1^4*f2^2+128*f0*f1^6*f2) * z1^2 * z2^2  +
```

```
(-96*f0^3*f1^3*f2^2+96*f0^2*f1^5*f2-96*f0^2*f1^3*f2^3+192*f0^2*f1^4*f2^2-192*f0*f1^6*f2
+96*f0*f1^5*f2^2+32*f0^4*f1*f2^3+64*f1^8-64*f0^3*f1^2*f2^3+32*f0^3*f1*f2^4-32*f0*f1^7
-32*f1^7*f2) * z1 * z2^3  +

(-96*f0^2*f1^4*f2^2-16*f0^4*f2^4-16*f1^8+64*f0*f1^6*f2+64*f0^3*f1^2*f2^3) * z2^4
```

The Likelihood at a marker locus

$$L(\theta, \nu) = z_0(\theta)^{n_0} \cdot z_1(\theta)^{n_1} \cdot z_2(\theta)^{n_2}$$

$$\nu = (p, f_{dd}, f_{Dd}, f_{DD})$$

One-Locus Log-Likelihood Ratio

By doing a Taylor expansion of the log-likelihood around $\theta_0 = \frac{1}{2}$ we find that we can write the log-likelihood ratio as

$$\begin{aligned} lr(\theta, \nu) &= 2(2\theta-1)^2(z_2 - z_0)\cdot(n_2 - n_0) \\ &\quad +(2\theta-1)^4(z_0 - z_1 + z_2)\cdot(n_0 - n_1 + n_2) \\ &\quad -2(2\theta-1)^4(z_2 - z_0)^2\cdot(n_0 + n_2) \\ &\quad +2(2\theta-1)^6(z_2 - z_0)(z_0 - z_1 + z_2)\cdot(n_2 - n_0) \\ &\quad -\frac{8}{3}(2\theta-1)^6(z_2 - z_0)^3\cdot(n_2 - n_0) \\ &\quad +\ldots \end{aligned}$$

A few important quantities

We define

$$\begin{aligned} w_L &= z_2 - z_0 \\ w_Q &= z_0 - z_1 + z_2 \end{aligned}$$

It is easy to show that

$$\begin{aligned} z_2(\theta) - z_0(\theta) &= (2\theta - 1)^2 (z_2 - z_0) \\ z_0(\theta) - z_1(\theta) + z_2(\theta) &= (2\theta - 1)^4 (z_0 - z_1 + z_2) \end{aligned}$$

A closer look at the log-likelihood ratio

Every term of the log-likelihood ratio can be written in terms of w_L and w_Q. The terms that appear are:

$$\begin{aligned}
&w_L \\
&w_L^2, \quad w_Q \\
&w_L^3, \quad w_L \cdot w_Q \\
&w_L^4, \quad w_L^2 \cdot w_Q, \quad w_Q^2 \\
&\dots
\end{aligned}$$

Notice:

$$w_L, w_Q \in [0, 1]$$

Two-Locus Linkage Analysis

- ◇ Identify new loci, that were not significant in a one locus analysis.
- ◇ Identify interactions between loci.
- ◇ Biological interaction vs. statistical interaction.

Two-Locus IBD probabilities

We now consider the joint IBD sharing at two loci,

$$z_{ij} = Pr(IBD = (i, j)|ASP)$$

It is convenient to display these 9 probabilities in a table:

z_{00}	z_{01}	z_{02}	z_{0+}
z_{10}	z_{11}	z_{12}	z_{1+}
z_{20}	z_{21}	z_{12}	z_{2+}
z_{+0}	z_{+1}	z_{+2}	1

Two-Locus Disease Models

We now have nine penetrances:

	bb	Bb	BB
aa	$f_{aa,bb}$	$f_{aa,Bb}$	$f_{aa,BB}$
Aa	$f_{Aa,bb}$	$f_{Aa,Bb}$	$f_{Aa,BB}$
AA	$f_{AA,bb}$	$f_{Aa,Bb}$	$f_{Aa,BB}$

Two locus disease models can be defined by picking your two favorite one-locus models and combining the penetrances:

Multiplicative model $f_{i_1j_1,i_2j_2} = x_{i_1j_1} \cdot y_{i_2j_2}$

Additive model $f_{i_1j_1,i_2j_2} = \min(x_{i_1j_1} + y_{i_2j_2}, 1)$

Heterogeneity model $f_{i_1j_1,i_2j_2} = x_{i_1j_1} + y_{i_2j_2} - x_{i_1j_1} \cdot y_{i_2j_2}$

The frequency of allele A is p_1 and that of B is p_2. The recombination fraction between gene i and marker i is θ_i, $i = 1, 2$.

Defining Polynomials

Multiplicative Recessive-Recessive

$z_{21}^2 - 4z_{20}z_{22}$	$z_{12}z_{21} - z_{11}z_{22}$	$z_{11}z_{21} - 4z_{10}z_{22}$
$z_{02}z_{21} - z_{01}z_{22}$	$z_{01}z_{21} - 4z_{00}z_{22}$	$z_{12}z_{20} - z_{10}z_{22}$
$z_{11}z_{20} - z_{10}z_{21}$	$z_{02}z_{20} - z_{00}z_{22}$	$z_{01}z_{20} - z_{00}z_{21}$
$z_{12}^2 - 4z_{02}z_{22}$	$z_{11}z_{12} - 4z_{01}z_{22}$	$z_{10}z_{12} - 4z_{00}z_{22}$
$z_{11}^2 - 16z_{00}z_{22}$	$z_{10}z_{11} - 4z_{00}z_{21}$	$z_{02}z_{11} - z_{01}z_{12}$
$z_{01}z_{11} - 4z_{00}z_{12}$	$z_{10}^2 - 4z_{00}z_{20}$	$z_{02}z_{10} - z_{00}z_{12}$
$z_{01}z_{10} - z_{00}z_{11}$	$z_{01}^2 - 4z_{00}z_{02}$	

Two-locus likelihood ratio

A Taylor expansion of the log-likelihood around $\theta_1 = \frac{1}{2}$ and $\theta_2 = \frac{1}{2}$.

$\Rightarrow$ The log-likelihood ratio can be written in terms of 8 quantities,

$$w_{L,0}, w_{0,L}, w_{Q,0}, w_{0,Q}, w_{L,L}, w_{L,Q}, w_{Q,L}, \quad \text{and} \quad w_{Q,Q}$$

that correspond to the one locus quantities w_L and w_Q.

Reminder

$$\begin{pmatrix} w_0 \\ w_L \\ w_Q \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \\ z_2 \end{pmatrix}$$

A linear transformation

$$\begin{pmatrix} w_{0,0} \\ w_{0,L} \\ w_{0,Q} \\ w_{L,0} \\ w_{L,L} \\ w_{L,Q} \\ w_{Q,0} \\ w_{Q,L} \\ w_{Q,Q} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}^{\otimes 2} \begin{pmatrix} z_{00} \\ z_{01} \\ z_{02} \\ z_{10} \\ z_{11} \\ z_{12} \\ z_{20} \\ z_{21} \\ z_{22} \end{pmatrix}$$

z_{00}	z_{01}	z_{02}
z_{10}	z_{11}	z_{12}
z_{20}	z_{21}	z_{12}

$w_{L,0}$

-1	-1	-1
0	0	0
1	1	1

$w_{0,L}$

-1	0	1
-1	0	1
-1	0	1

$w_{Q,0}$

1	1	1
-1	-1	-1
1	1	1

$w_{0,Q}$

1	-1	1
1	-1	1
1	-1	1

$w_{L,L}$

1	0	-1
0	0	0
-1	0	1

$w_{L,Q}$

-1	1	-1
0	0	0
1	-1	1

$w_{Q,L}$

-1	0	1
1	0	-1
-1	0	1

$w_{Q,Q}$

1	-1	1
-1	1	-1
1	-1	1

$$
\begin{aligned}
lr(\theta_1, \theta_2, \nu) \quad = \quad & 2w_1 n_1 + 2w_2 n_2 \\
& + w_3 n_3 + w_4 n_4 \\
& - 2w_1^2 (n_{0+} + n_{2+}) - 2w_2^2 (n_{+0} + n_{+2}) \\
& + 4(w_5 - w_1 w_2) n_5 \\
& + \frac{8}{3} w_1^3 n_1 - 2w_1 w_3 n_1 \\
& + \frac{8}{3} w_2^3 n_2 - 2w_2 w_4 n_2 \\
& + 2(w_6 - w_1 w_4) n_6 \\
& - 8w_1 (w_5 - w_1 w_2)(n_{02} - n_{00} + n_{22} - n_{20}) \\
& + 2(w_7 - w_2 w_3) n_7 \\
& - 8w_2 (w_5 - w_1 w_2)(n_{20} - n_{00} + n_{22} - n_{02}) + \ldots
\end{aligned}
$$

$$
\begin{aligned}
lr(\theta_1, \theta_2, \nu) \quad = \quad & \ldots \\
& +(w_8 - w_3 w_4) n_8 \\
& -4w_1^4(n_{0+} + n_{2+}) - \frac{1}{2} w_3^2 n_3 + 4w_1^2 w_3 (n_{0+} + n_{2+}) \\
& -4w_2^4(n_{+0} + n_{+2}) - \frac{1}{2} w_4^2 n_4 + 4w_2^2 w_4 (n_{+0} + n_{+2}) \\
& +16w_1^2(w_5 - w_1 w_2) n_5 - 4w_1(w_6 - w_1 w_4) n_5 - 4w_2(w_5 - w_1 w_2) n_5 \\
& +16w_2^2(w_5 - w_1 w_2) n_5 - 4w_2(w_7 - w_2 w_3) n_5 - 4w_2(w_5 - w_1 w_2) n_5 \\
& +24w_1 w_2 (w_5 - w_1 w_2)(n_{00} + n_{02} + n_{20} + n_{22}) \\
& +4w_1(w_7 - w_2 w_3) n_7 \\
& +4w_2(w_6 - w_1 w_4) n_6
\end{aligned}
$$

Two-Locus Tests

The two locus log-likelihood ratio can be written as the sum of:

> the log-likelihood ratio for locus 1,
>
> the log-likelihood ratio for locus 2 and
>
> interaction terms.

Question of interest: How big are the interaction terms? I.e. how much do we gain by doing a two locus analysis over a one locus analysis?

The terms in the log-likelihood ratio

$$w_{L,0}$$

$$w_{0,L}$$

$$w_{Q,0}$$

$$w_{0,Q}$$

$$w_{L,L} - w_{L,0}w_{0,L}$$

$$w_{L,Q} - w_{L,0}w_{0,Q}$$

$$w_{Q,L} - w_{Q,0}w_{0,L}$$

$$w_{Q,Q} - w_{Q,0}w_{0,Q}$$

Sample correlations

Consider just one affected sib-pair.

Let N_{ij} be indicators for their joint IBD sharing.

We use the same notation as for the w's, $N_{L,0} = N_{2+} - N_{0+}$ etc.,

$$Corr(N_{L,0}, N_{0,L}) = \frac{w_{L,L} - w_{L,0}w_{0,L}}{\sqrt{(w_{Q,0}+1)/2 + w_{L,0}^2}\sqrt{(w_{0,Q}+1)/2 + w_{0,L}^2}}$$

$$Corr(N_{L,0}, N_{0,Q}) = \frac{w_{L,Q} - w_{L,0}w_{0,Q}}{\sqrt{(w_{Q,0}+1)/2 + w_{L,0}^2}\sqrt{1 - w_{0,Q}^2}}$$

$$Corr(N_{Q,0}, N_{0,L}) = \frac{w_{Q,L} - w_{Q,0}w_{0,L}}{\sqrt{1 - w_{Q,0}^2}\sqrt{(w_{0,Q}+1)/2 + w_{0,L}^2}}$$

$$Corr(N_{Q,0}, N_{0,Q}) = \frac{w_{Q,Q} - w_{Q,0}w_{0,Q}}{\sqrt{1 - w_{Q,0}^2}\sqrt{1 - w_{0,Q}^2}}$$

Statistical Independence

We say that there are no interactions in the table of z_{ij}'s if

$$z_{ij} = z_{i+} z_{+j}$$

This implies that

$$\begin{aligned} w_{L,L} - w_{L,0} \cdot w_{0,L} &= 0 \\ w_{L,Q} - w_{L,0} \cdot w_{0,Q} &= 0 \\ w_{Q,L} - w_{Q,0} \cdot w_{L,0} &= 0 \\ w_{Q,Q} - w_{Q,0} \cdot w_{0,Q} &= 0 \end{aligned}$$

So all the interaction terms in the log-likelihood are 0.

Note: If the penetrances factorize, i.e. $f_{ij} = a_i b_j$ where a_i and b_j are the single locus penetrances then it follows that $z_{ij} = z_{i+} z_{+j}$.

Defining Polynomials

A two locus multiplicative Recessive-Recessive model

$$
\begin{aligned}
w_{L,0}^2 - w_{Q,0} &= 0 \\
w_{0,L}^2 - w_{0,Q} &= 0 \\
w_{L,L} - w_{L,0} \cdot w_{0,L} &= 0 \\
w_{L,Q} - w_{L,0} \cdot w_{0,Q} &= 0 \\
w_{Q,L} - w_{Q,0} \cdot w_{L,0} &= 0 \\
w_{Q,Q} - w_{Q,0} \cdot w_{0,Q} &= 0
\end{aligned}
$$

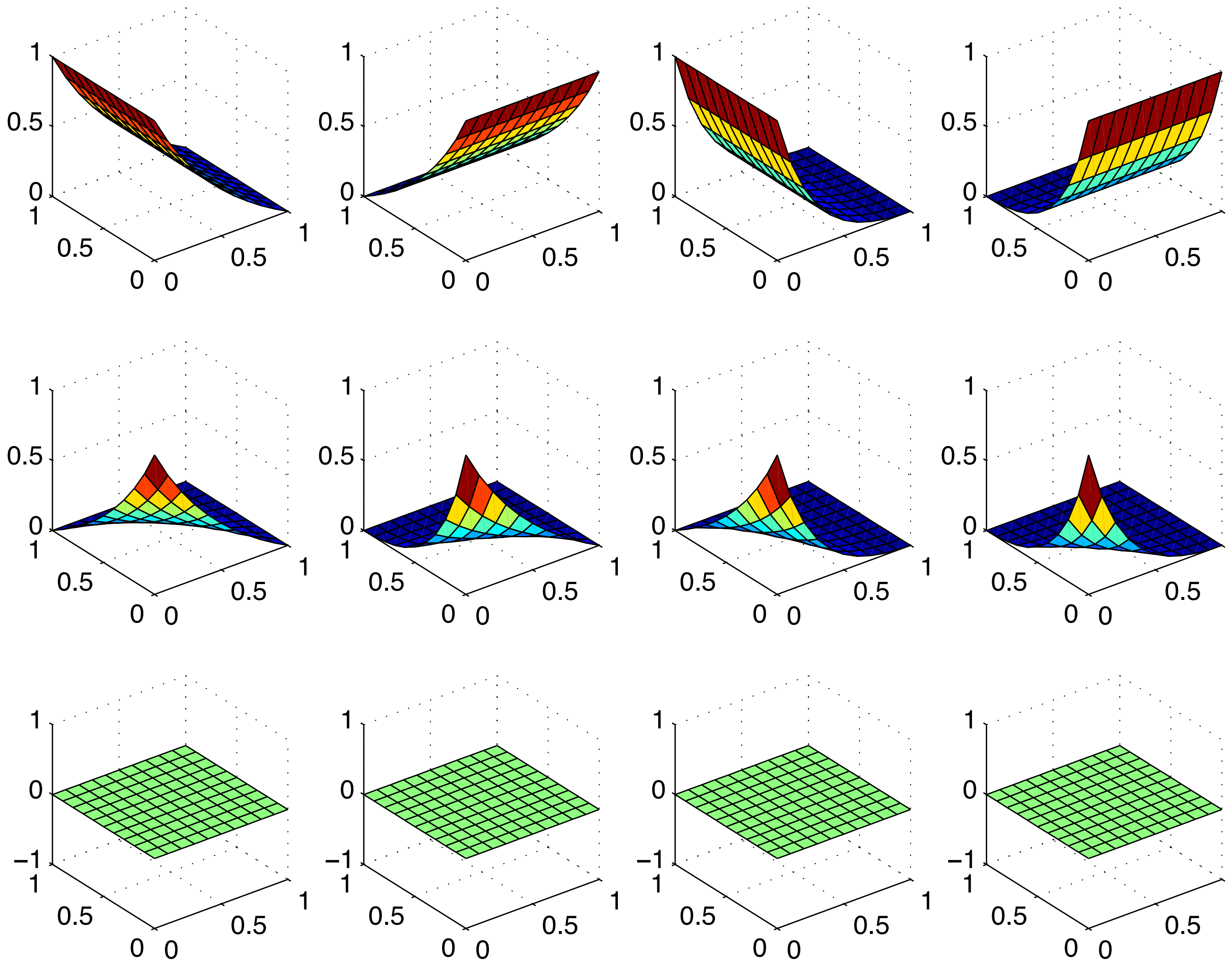
Plots

$w_{L,0}$	$w_{0,L}$	$w_{Q,0}$	$w_{0,Q}$
$w_{L,L}$	$w_{L,Q}$	$w_{Q,L}$	$w_{Q,Q}$
$w_{L,L} - w_{L,0}w_{0,L}$	$w_{L,Q} - w_{L,0}w_{0,Q}$	$w_{Q,L} - w_{Q,0}w_{0,L}$	$w_{Q,Q} - w_{Q,0}w_{0,Q}$

Multiplicative Recessive-Recessive

Penetrance table

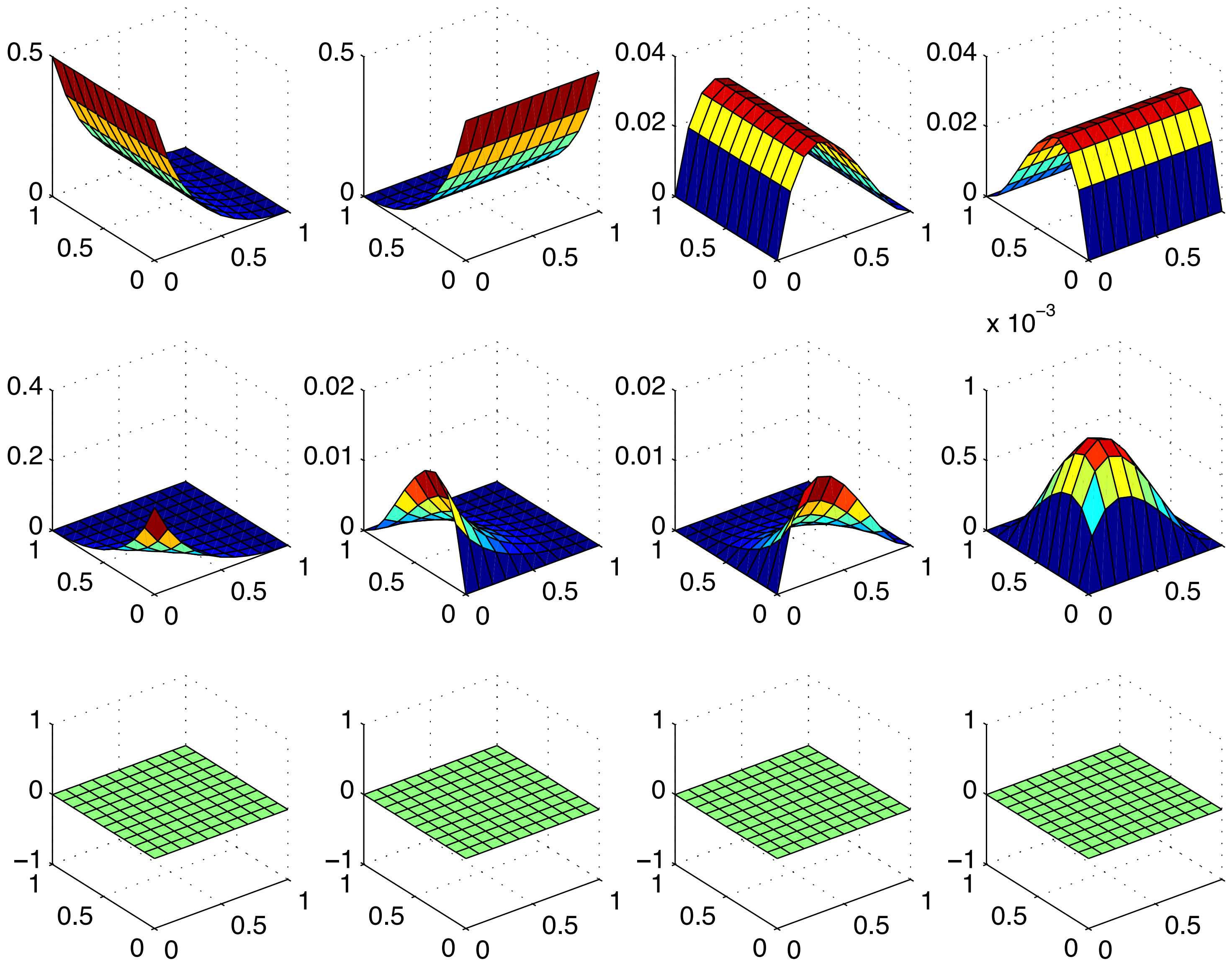
	bb	Bb	BB
aa	0	0	0
Aa	0	0	0
AA	0	0	1



Multiplicative Dominant-Dominant

Penetrance table

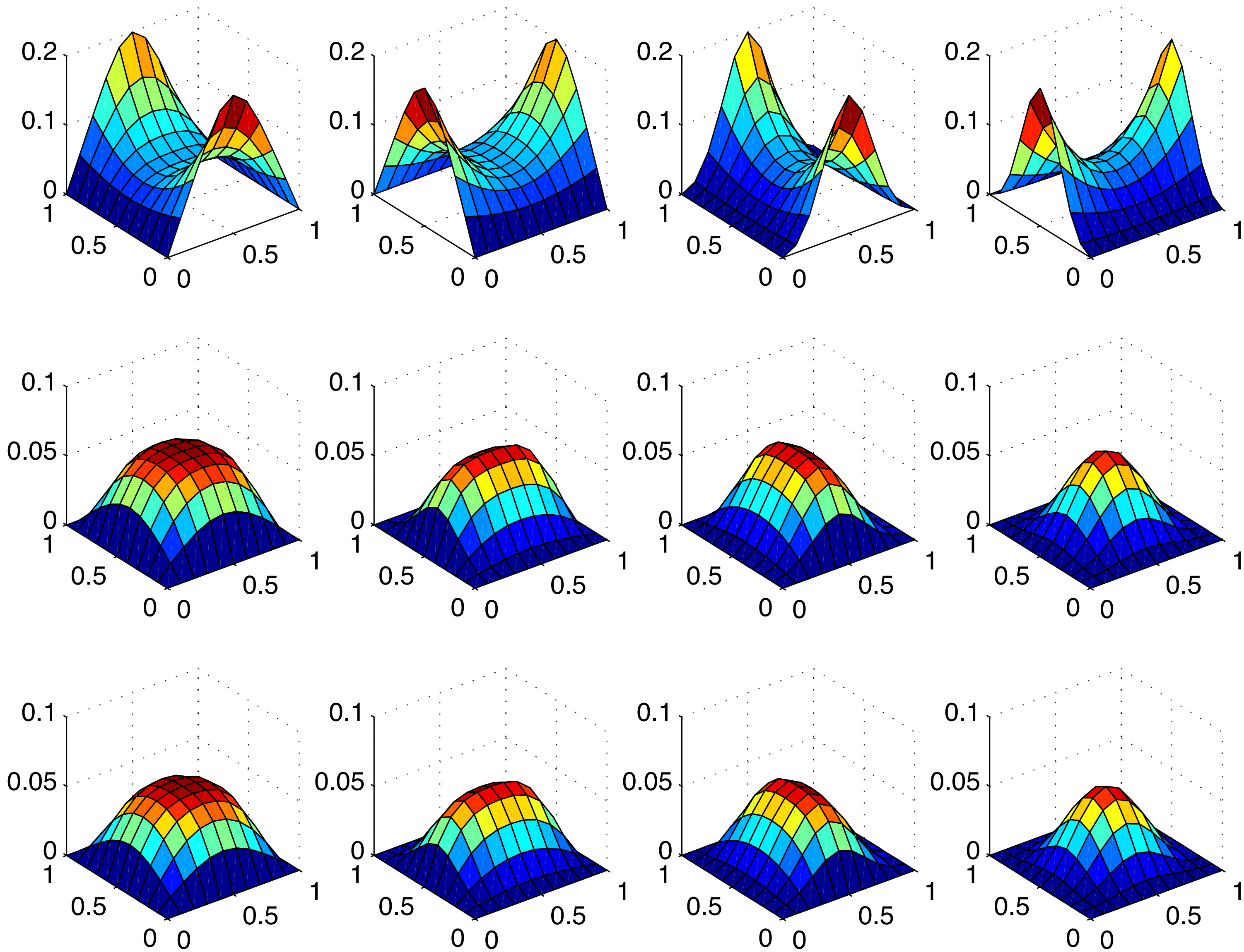
	bb	Bb	BB
aa	0	0	0
Aa	0	1	1
AA	0	1	1


x 10^−3

Butterfly model

Penetrance table

	bb	Bb	BB
aa	1	0	1
Aa	0	1	0
AA	1	0	1


0.2
0.1
0
1
0.5
0
0
0.5
1
0.1
0.05

Variance Components

$$\begin{pmatrix} w_{L,0} \\ w_{0,L} \\ w_{Q,0} \\ w_{0,Q} \\ w_{L,L} \\ w_{L,Q} \\ w_{Q,L} \\ w_{Q,Q} \end{pmatrix} = \frac{1}{16KK_s} \left(\begin{array}{cccc|cccc} 4 & 0 & 4 & 0 & 2 & 1 & 2 & 1 \\ & 4 & 0 & 4 & 2 & 2 & 1 & 1 \\ & & 4 & 0 & 0 & 0 & 2 & 1 \\ & & & 4 & 0 & 2 & 0 & 1 \\ \hline & & & & 1 & 1 & 1 & 1 \\ & & & & & 1 & 0 & 1 \\ & & & & & & 1 & 1 \\ & & & & & & & 1 \end{array} \right) \begin{pmatrix} \sigma^2_{A_1} \\ \sigma^2_{A_2} \\ \sigma^2_{D_1} \\ \sigma^2_{D_2} \\ \sigma^2_{A_1 A_2} \\ \sigma^2_{A_1 D_2} \\ \sigma^2_{D_1 A_2} \\ \sigma^2_{D_1 D_2} \end{pmatrix}$$

Summary

◇ Change of co-ordinate system, from z's to w's.

◇ The two locus interaction terms in the log-likelihood ratio can all be written in terms of $w_{i,j} - w_{i,0}w_{0,j}$.

◇ The interaction terms can be estimated from four sample correlations.

◇ For most models the increase in the LR, beyond two one locus analyses, is modest.

Acknowledgements

Terry Speed

Bernd Sturmfels

Sandrine Dudoit

Olof Bengtsson