Chapter 5

Models for Meiosis

5.1 The meiosis process

In section 4.1, we introduced recombination as the process of crossing over between the two homologous parental chromosomes in the formation of an offspring gamete, and we have considered multilocus segregation probabilities under the assumption of no interference (section 4.7). In order to develop better models of multilocus segregation, it is necessary to consider the processes of *mitosis* and *meiosis* in greater detail. Mitosis is the normal process of cell division during somatic growth: meiosis is the process of gamete formation. Both processes involve chromosome duplication and separation, but only meiosis involves recombination. A chromosome is a doubled strand of helical DNA, with complementary bases on the two strands. Chromosomes of the shape often depicted in texts, or seen in an amniocentesis photograph, exist only just prior to mitosis or meiosis. These are actually doubled chromosomes. Each chromosome is thus two double strands of DNA. Each double-strand is known as a *chromatid*: the two chromatids of a single duplicate chromosome are known as *sister* chromatids. In the pair of chromosomes just prior to mitosis or meiosis, there are thus four chromatids, or eight strands of DNA in total. In our modeling here, we consider the four chromatids, or the *chromatid tetrad*, rather than all eight DNA strands.

Just after the previous mitotic division, each chromosome exists as a concentrated double-strand of DNA in the nucleus of the cell (Figure 5.1(a)). In the next stage, *interphase*, the chromosomes elongate (Figure 5.1(b)), and duplicate; at this stage the length of DNA in the nucleus of a cell is 2 meters. The DNA then re-concentrates to form the chromatid tetrad (Figure 5.1(c)). In mitosis, each chromosome divides to give two daughter cells (Figure 5.1(d)), each with a nucleus with the identical chromosome complement as the parent cell nucleus (Figure 5.1(a)). In the first meiotic division, however, one of each homologous pair of chromosomes must go to each daughter cell. In order to achieve this, the pair of chromosomes must become tightly aligned, and in so doing *chiasmata* occur, resulting in an exchange of DNA between two non-sister chromatids (Figure 5.1(e)).

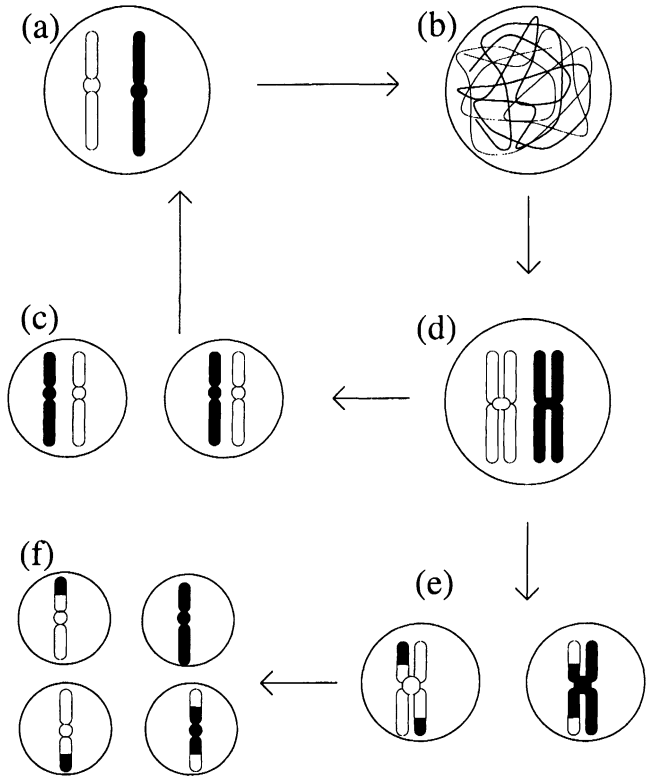

FIGURE 5.1. *The processes of mitosis and meiosis, shown for a single pair of homologous chromosomes in the nucleus of a cell of a diploid organism. See text for details*

The chromosomes separate; each daughter cell nucleus now contains only 50% of the DNA of the parent cell, but still in duplicate chromatid form (Figure 5.1(f)). Finally in the second meiotic division, in a process analogous to mitosis, these chromosomes divide, providing potential gamete cells (Figure 5.1(g)). Each potential gamete now contains 50% of the parental DNA, in the haploid form of one chromosome from each chromosome pair.

The crossover process is shown in more detail in Figure 5.2. Figure 5.2(a) shows the tetrad on which, in this example, two chiasmata are formed, and Figure 5.2(b) shows the four resulting gamete chromosomes. In mammalian organisms, for male meioses all four become gametes (sperm), while in female meioses three are discarded and one becomes a gamete (egg cell). However, only for certain non-mammalian species (such as fungi), or by carefully designed experiments (Hulten et al., 1990), is it possible to retrieve the four sperm from a given meiosis. In the analysis of data on an offspring individual, we observe only one paternal and one maternal meiotic product.

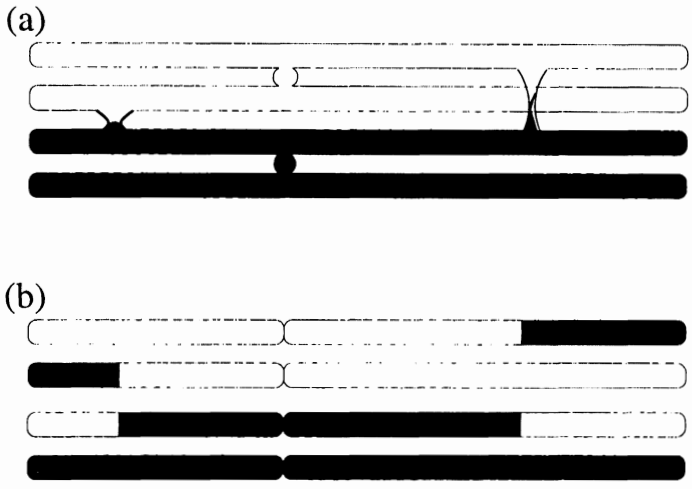

FIGURE 5.2. *The formation of chiasmata, and the crossovers resulting in the chromosomes of the four offspring gametes. The crossovers occurring are the same as in Figure 5.1(e)*

5.2 From chromatids to crossovers

Instead of modeling the crossover locations in a gamete (section 4.1), we now consider the occurrence of chiasmata locations at which crossovers between non-sister chromatids occur. Models for chiasmata formation are known as *four-strand* models, since the four chromatids are considered. Since each chiasma involves one paternal and one maternal chromatid, (paternal and maternal referring to the grandparental origins of the two homologous parental chromosomes as in equation (1.2)), each chiasma exists as a crossover in a resulting gamete with marginal probability $\frac{1}{2}$. Recall that the definition of genetic distance, provides for an expected one crossover per Morgan (section 4.1): this corresponds to an expectation of two chiasmata per Morgan, or one per 50 centiMorgans (cM).

Where only one meiotic product is observed, obtaining evidence for chromatid interference is practically impossible (Zhao et al., 1995) (but see also section 5.5). It is therefore often assumed that there is no *chromatid interference*: that is, that each chiasma involves two randomly chosen non-sister chromatids, independently of other the chromatids involved in other chiasmata. In this case, each chiasma results in a crossover in a given gamete, independently with probability $\frac{1}{2}$. Or the crossover process is just a thinned (probability= $\frac{1}{2}$) version of the chiasma process. Since a thinned Poisson process is also a Poisson process, this has no impact on the Haldane (1919) no-interference model. The chiasma process is Poisson, rate 2 per Morgan. The crossover process is Poisson, rate 1 per Morgan.

More generally, in a given chromosome interval of genetic length d, suppose there are $N(d)$ chiasmata, making now no assumptions about the probability distribution of N. If $N(d) = 0$, there are no chiasmata, no crossovers, and hence no recombination. For any non-zero value n of $N(d)$, in the absence of chromatid

interference, the probability of an odd number of crossovers is 1/2. (This is left as an exercise to the reader: it may be easier to think about tossing a fair coin n times, and the probability of an odd number of "heads".) Thus we have the formula of Mather (1938) for the recombination probability $\rho(d)$ at genetic distance d:

$$\rho(d) = \frac{1}{2}\Pr(N(d) > 0) = \frac{1}{2}(1 - \Pr(N(d) = 0)). \tag{5.1}$$

The only assumption here is the absence of chromatid interference: under this assumption $\rho(d)$ is an increasing function of d, and is bounded above by $\frac{1}{2}$. Note also that under Haldane's model $\Pr(N(d) = 0) = \exp(-2d)$, and Mather's formula applies (see equation (4.2)).

5.3 From chiasmata to recombination patterns

There is a multilocus version of Mather's formula (5.1). As in section 4.7, consider a chromosome with L ordered loci, $1, \ldots, L$, and label the intervals $I_1, \ldots, I_{L-1}$ and let $R_j = 1$ if a gamete is recombinant on interval I_j, and $R_j = 0$ otherwise ($j = 1, \ldots, L-1$). The recombination pattern is a function of the meiosis indicators $S_{i,j}$ for the given meiosis i, and provides a simpler representation for the current discussion:

$$\begin{aligned} R_j &= 0 \quad \text{if } S_{i,j} = S_{i,j+1} \\ R_j &= 1 \quad \text{if } S_{i,j} \neq S_{i,j+1} \end{aligned}$$

for $j = 1, \ldots, L-1$.

Now also let the (random) number of chiasmata in the intervals, in a meiosis, be $N_1, \ldots, N_{L-1}$. Let $C_j = 0$ if $N_j = 0$, and $C_j = 1$ otherwise ($j = 1, \ldots, L-1$): C_j is an indicator of presence of chiasmata in interval I_j. If $C_j = 0$, then $R_j = 0$. If $C_j = 1$, then $\Pr(R_j = 1) = \Pr(R_j = 0) = \frac{1}{2}$. In the absence of chromatid interference, the R_j are conditionally independent given C_j. Thus

$$\begin{aligned} \Pr((R_1, \ldots, R_{L-1}) = \mathbf{r}) &= \sum_{\mathbf{c} \geq \mathbf{r}} (\frac{1}{2})^{|\mathbf{c}|} \Pr((C_1, \ldots, C_{L-1}) = \mathbf{c}) \\ &= (\frac{1}{2})^{L-1} \sum_{\mathbf{c} \geq \mathbf{r}} 2^{|\mathbf{1}-\mathbf{c}|} \Pr((C_1, \ldots, C_{L-1}) = \mathbf{c}) \end{aligned} \tag{5.2}$$

where $|\mathbf{c}| = \sum_1^{L-1} c_j$ is the number of unit indicators in $\mathbf{c}$, and $\mathbf{1}$ is a vector of ones. This equation is (in essence) due to Weinstein (1936). Karlin and Liberman (1979) give a version in terms of the meiosis indicators rather than the recombination indicators. A recent discussion, using slightly different notation, is given by Speed (1996).

The estimation of chiasmata presence and absence patterns from recombination data provides another example of use of the EM algorithm. Consider again equation (5.2), and the estimation of patterns of chiasmata presence and absence,

from a sample of n completely observed patterns, $\mathbf{r}$, of recombination and non-recombination. An unconstrained estimate of $\Pr(\mathbf{R} = \mathbf{r})$ is $n(\mathbf{r})/n$ where $n(\mathbf{r})$ is the number of meioses exhibiting recombination patterns $\mathbf{r}$. However, if the equation

$$n(\mathbf{r}) = n \sum_{\mathbf{c} \geq \mathbf{r}} (\frac{1}{2})^{|\mathbf{c}|} \Pr((C_1, \ldots, C_{L-1}) = \mathbf{c}) \tag{5.3}$$

is inverted, negative values of $\Pr(\mathbf{C} = \mathbf{c})$ may result. An EM algorithm (section 2.4) avoids this, providing estimates of the probabilities of the underlying chiasmata presence/absence patterns, $q(\mathbf{c}) = \Pr((C_1, \ldots, C_{L-1}) = \mathbf{c})$, subject only to the constraint of no chromatid interference. In fact, this EM algorithm is very similar to that of section 4.2. There a phenotypic observation was partitioned in expectation among the possible multilocus genotypes (pairs of haplotypes) providing that phenotype. Here observation of a recombination pattern is subdivided among the chiasmata presence/absence patterns that could give rise to the recombination pattern:

$$\begin{aligned} \Pr((C_1, \ldots, C_{L-1}) = \mathbf{c} \mid \mathbf{r}) &= \frac{(\frac{1}{2})^{|\mathbf{c}|}\, q(\mathbf{c})}{\sum_{\mathbf{c}^* \geq \mathbf{r}} (\frac{1}{2})^{|\mathbf{c}^*|}\, q(\mathbf{c}^*)} \quad \text{if } \mathbf{c} \geq \mathbf{r} \\ &= 0 \qquad\qquad\qquad\qquad \text{otherwise.} \end{aligned}$$

Thus, given current estimates $q(\mathbf{c})$ and the data counts $n(\mathbf{r})$, the conditional expected number of meioses exhibiting chiasmata pattern $\mathbf{c}$ is

$$\sum_{\mathbf{r} \leq \mathbf{c}} n(\mathbf{r}) \frac{(\frac{1}{2})^{|\mathbf{c}|}\, q(\mathbf{c})}{\sum_{\mathbf{c}^* \geq \mathbf{r}} (\frac{1}{2})^{|\mathbf{c}^*|}\, q(\mathbf{c}^*)}$$

and the new estimate is simply n^{-1} times this expected number. This EM algorithm, although very simply implemented, has poor convergence if there are many loci, or very tightly linked loci, since then many patterns $\mathbf{c}$ do not occur in the sample. Moreover, the resulting constrained MLEs differ from the inversion of (5.3) only when some $\Pr(\mathbf{C} = \mathbf{c})$ have MLE 0. In this case, unfortunately, convergence of the EM algorithm can be very slow. However, again as in the case of section 4.2, some frequencies $q(\mathbf{c})$ may be constrained to zero, and estimation of other chiasmata pattern frequencies continued in the subspace.

5.4 The chiasmata avoidance process

The vector $(C_1, \ldots, C_{L-1})$, specifies the avoidance and non-avoidance probabilities of the *chiasma process* on intervals of the chromosome. It is slightly neater, although of course equivalent, to express $\Pr((R_1, \ldots, R_{L-1}) = \mathbf{r})$ (or the probability of gametic types $\Pr(S_{i,\bullet})$) in terms of the avoidance probabilities alone, as in Mather's formula (5.1). We specify a subset $\mathcal{T}$ of the intervals $\{I_1, ..., I_{L-1}\}$ as follows. Let $t_j = 1$ if $I_j \in \mathcal{T}$, and $t_j = 0$ otherwise. Let $\phi_{\mathbf{t}}$ be the probability of no chiasmata

in $\mathcal{T}$. The set of $\phi_{\mathbf{t}}$, for all binary vectors $\mathbf{t}$ length $(L-1)$, is the set of avoidance probabilities of the chiasma process. If $t_j = 1$ there are no chiasmata in I_j, but if $t_j = 0$ the presence/absence of chiasmata in I_j is unspecified. There is thus a one-one relationship between the avoidance probabilities $\phi_{\mathbf{t}}$ and the presence/absence probabilities $\Pr(C_1, \ldots, C_{L-1})$:

$$\begin{aligned} \phi_{\mathbf{t}} &= \Pr(\text{no chiasmata in } \mathcal{T}) \\ &= \sum_{\mathbf{c} \leq (\mathbf{1}-\mathbf{t})} \Pr((C_1, \ldots, C_{L-1}) = \mathbf{c}). \end{aligned} \tag{5.4}$$

Lange (1997) derives an expression

$$\Pr((R_1, \ldots, R_{L-1}) = \mathbf{r}) = (\frac{1}{2})^{L-1} \sum_{\mathbf{t}} (-1)^{<\mathbf{r},\mathbf{t}>} \phi_{\mathbf{t}} \tag{5.5}$$

by a different method, again with notation differing slightly from ours. (Here, $<\mathbf{r},\mathbf{t}>$ is the inner product $\sum_j r_j t_j$.) Rather than deriving this equation directly, we use equation (5.4) to show that (5.2) and (5.5) are equivalent. Substituting (5.4) into (5.5) we obtain

$$\begin{aligned} \sum_{\mathbf{t}} (-1)^{<\mathbf{r},\mathbf{t}>} \phi_{\mathbf{t}} &= \sum_{\mathbf{t}} (-1)^{<\mathbf{r},\mathbf{t}>} \left(\sum_{\mathbf{c} \leq (\mathbf{1}-\mathbf{t})} \Pr(\mathbf{C} = \mathbf{c}) \right) \\ &= \sum_{\mathbf{c}} \left(\sum_{\mathbf{t} \leq (\mathbf{1}-\mathbf{c})} (-1)^{<\mathbf{r},\mathbf{t}>} \right) \Pr(\mathbf{C} = \mathbf{c}). \end{aligned}$$

Equating coefficients of $\Pr(\mathbf{C} = \mathbf{c})$ from (5.2), to complete the proof we need only show that for each $\mathbf{r}$ and $\mathbf{c}$

$$2^{|\mathbf{1}-\mathbf{c}|} I\{\mathbf{c} \geq \mathbf{r}\} = \sum_{\mathbf{t} \leq (\mathbf{1}-\mathbf{c})} (-1)^{<\mathbf{r},\mathbf{t}>}$$

where $I\{\mathbf{c} \geq \mathbf{r}\} = 1$ if $\mathbf{c} \geq \mathbf{r}$, and 0 otherwise. Consider first the case $\mathbf{c} \geq \mathbf{r}$. Then $r_j = 1 \Rightarrow c_j = 1 \Rightarrow t_j = 0$, so $<\mathbf{r},\mathbf{t}>= 0$ and we sum terms $(+1)$ over $2^{|\mathbf{1}-\mathbf{c}|}$ values of $\mathbf{t}$, confirming this case. Now consider any other $\mathbf{c}$, and consider any one component j for which $r_j = 1$ but $c_j = 0$. Thus $1 - c_j = 1$, and we sum over $t_j = 0$ and $t_j = 1$. For each set of values of the other $t_{j'}$, the two values of t_j give opposite signs to $(-1)^{<\mathbf{r},\mathbf{t}>}$. The coefficients cancel, and the overall coefficient is 0, as required. This completes the proof.

Given a model which determines either $\Pr(\mathbf{C} = \mathbf{c})$ or $\phi_{\mathbf{t}}$, exact computation of probabilities $\Pr(\mathbf{R} = \mathbf{r})$ of all patterns of recombination and non-recombination in a set of $L-1$ marker intervals is practical for L up to about 12. Two methods of likelihood evaluation under interference have been proposed: both rely on efficient computation of these probabilities. Weeks et al. (1993) provides an approach for models of count interference (see section 5.6), while Lin and Speed (1996) provides a method for the renewal process chi-square models of position interference

(section 5.7). For a fixed marker map, it is feasible to precompute and store these probabilities for up to about 12 markers ($2^{11} = 2048$). However, using these probabilities in any exact computation of a likelihood on a pedigree usually entails highly computationally intensive procedures, further limiting L and/or the pedigree sizes and structures that can be considered.

5.5 Chromatid interference

Where only one of the four gametic products of meiosis can be observed, it is hard to find evidence for chromatid interference. However, the non-negativity of probabilities $P(\mathbf{C} = \mathbf{c})$ in equation (5.2) does impose constraints on feasible recombination pattern probabilities $P(\mathbf{R} = \mathbf{r})$. Conversely, observed frequencies of patterns of recombination can provide evidence for the existence of chromatid interference. We consider now one specific constraint implied by Mather's formula, whose violation may provide evidence for chromatid interference (Fisher, 1948). Mather's formula (5.1) implies that recombination probabilities are an increasing function of genetic distance, bounded above by $\frac{1}{2}$. Under chromatid interference this is no longer so. Consider, in particular, the case of *complete positive chromatid interference*: in that case, successive chiasmata involve alternating disjoint pairs of non-sister chromatids. Then the recombination probability at genetic distance d is

$$\rho(d) \;=\; \frac{1}{2}\Pr(N(d) \text{ odd}) \;+\; \Pr(N(d) \text{ even but not divisible by } 4).$$

In the case when the chiasma process is a Poisson process rate 2, this becomes

$$\begin{aligned}\rho(d) &= \frac{1}{2}\exp(-2d)\left(\sum_{k=0}^{\infty}\left(\frac{(2d)^{2k+1}}{(2k+1)!} \;+\; 2\,\frac{(2d)^{4k+2}}{(4k+2)!}\right)\right)\\ &= \frac{1}{2}(1-\exp(-2d)\,\cos(2d)).\end{aligned}$$

In this case, $\rho(d)$ is greater than $\frac{1}{2}$ at certain distances, and is not monotone. Fisher (1948) discusses possible evidence for $\rho(d) > \frac{1}{2}$ in the case of the pseudoautosomal region of the mammalian sex chromosomes in mice; Weinstein provides an interesting contribution to the discussion.

Another possibility is *complete negative chromatid interference*: in this case every chiasma on the tetrad involved the same pair of chromatids. Then half the gametes would show no recombination, but in the other potential two gametes from a meiosis every chiasma results in a crossover. Again, when chiasmata occur as a Poisson process rate 2,

$$\rho(d) \;=\; \frac{1}{2}\Pr(N(d) \text{ odd}) \;=\; \frac{1}{4}(1-\exp(-4d)).$$

Note that when d is small, $\rho(d) \approx d$, as usual. However, at large genetic distances, only one half of the gametes show independent segregation, the other half apparently showing tight linkage. With multilocus data, such an extreme pattern of

recombination would be detectable. A less extreme pattern might simply be thought to be due to heterogeneity of recombination among meioses. Chromatid interference is very much confounded both with other forms of interference, and interference in general may be confounded with heterogeneity in recombination. For the remainder of this chapter we consider only models with no chromatid interference.

To ensure a biologically feasible interference model, a model of chiasma formation in the chromatid tetrad at meiosis is desirable. Under the no-interference model (Haldane, 1919), chiasmata, and hence crossovers, arise as a Poisson process; the count on a chromosome arm has a Poisson distribution, and conditionally on the count their positions are independently and uniformly distributed (all distributions being in terms of genetic, not physical, distance). Thus, in the absence of chromatid interference, there are, broadly, two classes of interference model: count interference and position interference.

5.6 Count-location models for chiasmata

In a count-location model, the count of chiasmata on a chromosome arm is no longer necessarily Poisson, but conditional upon the count, they are independently and uniformly distributed. In such models

$$\phi_{\mathbf{t}} \quad = \quad \phi(< \mathbf{t}, \mathbf{d} >)$$

where d_j is the genetic length of interval I_j. That is, the chiasma avoidance function depends only on the total length of chromosome avoided. Such models have been considered by Liberman and Karlin (1984), who call the corresponding map functions $\rho(d)$ *multilocus feasible.*

Suppose that the probability mass function of the total number of chiasmata N on a chromosome arm length Λ Morgans has probability generating function $g_N(\cdot)$. Then, given $N = n$, the probability of no chiasmata in length d is $(1 - d/\Lambda)^n$, and

$$\phi(d) \quad = \quad \sum_{n=0}^{\infty} \Pr(N = n)(1 - \frac{d}{\Lambda})^n \quad = \quad g_N(1 - d/\Lambda) \tag{5.6}$$

with corresponding map function, from Mather's formula,

$$\rho(d) \quad = \quad \frac{1}{2}(1 - \phi(d)) \quad = \quad \frac{1}{2}(1 - g_N(1 - d/\Lambda)). \tag{5.7}$$

Note that the expected number of chiasmata N in length Λ of chromosome is, by the definition of genetic length, 2Λ.

Consider now some simple examples:
(1) Suppose N has a Poisson distribution with mean 2Λ: $N \sim \mathcal{P}(2\Lambda)$
Then $g_N(w) = \mathrm{E}(w^N) = \exp(2\Lambda(w - 1))$ and from equation (5.6),

$$\phi(d) \quad = \quad \exp(2\Lambda(1 - \frac{d}{\Lambda} - 1)) \quad = \quad \exp(-2d)$$

and from equation (5.7) we have again the no-interference equation (4.2).

(2) Another tractable count-location model is given by assuming a fixed maximum number K of chiasmata on a chromosome, and that $N \sim \mathcal{B}(K, \frac{1}{2})$, with $2\Lambda = \mathrm{E}(N) = K/2$.
Then $g_N(w) = \mathrm{E}(w^N) = (\frac{1}{2}(1+z))^K$ and from equation (5.6),

$$\phi(d) \;=\; (\frac{1}{2}(2 - \frac{d}{\Lambda}))^K \;=\; (1 - \frac{d}{2\Lambda})^{4\Lambda}.$$

For large chromosomes, there is little interference: $\phi(d)$ becomes close to the non-interference value $\exp(-2d)$. On small chromosomes there is stronger interference. For example, if $\Lambda = \frac{1}{2}$, $\phi(d) = (1-d)^2$, $\rho(d) = d(1 - \frac{1}{2}d)$; the avoidance probability is smaller, and the recombination probability larger, than in the absence of interference.

(3) It appears to be be a biological reality, that for correct division of the chromosomes in the first meiotic division (Figure 5.1(d) to Figure 5.1(e)), each chromosome pair should have at least one chiasma. Note that under any such model $N \geq 1$ so that $\Lambda = \frac{1}{2}\mathrm{E}(N) \geq \frac{1}{2}$; in fact, even the smallest human autosomes have genetic length estimates just over 0.5 Morgans. One example of a model which incorporates this restriction is the truncated Poisson model, in which N has a Poisson distribution ($N \sim \mathcal{P}(\alpha)$) conditioned on $N \geq 1$. Then $2\Lambda = \mathrm{E}(N) = \alpha/(1 - \exp(-\alpha))$, and Λ is an increasing function of α, increasing from $\frac{1}{2}$ when $\alpha = 0$. Then

$$g_N(w) \;=\; \frac{\exp(\alpha(w-1)) - \exp(-\alpha)}{1 - \exp(-\alpha)} \quad \text{and} \quad \phi(d) \;=\; \frac{\exp(\alpha(1 - \frac{d}{\Lambda})) - 1}{\exp(\alpha) - 1}.$$

(4) An alternative model incorporating the restriction $N \geq 1$ is that due to Sturt (1976), in which N has a shifted, rather than truncated, Poisson distribution: $(N-1) \sim \mathcal{P}(2\Lambda - 1)$. Then

$$g_N(w) \;=\; w\exp((2\Lambda - 1)(w-1)) \quad \text{and} \quad \phi(d) \;=\; (1 - \frac{d}{\Lambda})\exp(-(2\Lambda - 1)\frac{d}{\Lambda}).$$

The Sturt model has been found to fit existing data well (Weeks et al., 1993).

All the models (2),(3) and (4) are close to the Haldane model on large chromosomes, but show different departures on small chromosomes. It is an unfortunate feature of count-location models that the recombination probability at genetic distance d is determined by the length of the chromosome and the distribution of N on the entire chromosome.

5.7 Renewal process models of chiasma formation

Although count-location models are convenient, mathematically, it is implausible that, given N, chiasmata are independently located. In particular, the consequence

that the chiasma avoidance function depends only on total length avoided is unrealistic. Consider two intervals, lengths d_1 and d_3 separated by an interval length d_2. Then, for a count location model, the probability

$$\phi_{(1,0,1)} = P(C_1 = 0, C_3 = 0) = \phi(d_1 + d_3)$$

and is independent of d_2. Position interference models allow for more general meiotic processes; we will consider only those where chiasmata arise as a stationary renewal process (Speed, 1996; Lange, 1997). This imposes certain restrictions on the map function $\rho(d)$, which are discussed by Speed (1996); subject to these restrictions, the renewal density is $-\rho''(d)$.

We consider briefly some examples: more details are given by Speed (1996) and references therein.

(1) Suppose chiasmata occur along the tetrad bundle as a Poisson process, rate 2, so that the interarrival time distribution is exponential with mean $\frac{1}{2}$, and has probability density function $2\exp(-2d)$. Integrating twice, and imposing the conditions $\rho(0) = 0$, and $\rho'(0) = 1$, we obtain again equation (4.2), confirming this interpretation of the no-interference model.

(2) Kosambi (1944) proposed a map function

$$\rho(d) = \frac{1}{2}\tanh(2d) = \frac{1}{2}\left(\frac{\exp(4d) - 1}{\exp(4d) + 1}\right)$$

which satisfies the conditions detailed by Speed (1996) and results in a renewal density

$$16\frac{(\exp(2d) - \exp(-2d))}{(\exp(2d) + \exp(-2d))^3}.$$

Although this map function is not *multilocus feasible* in the sense of Liberman and Karlin (1984), it has a valid interpretation as the result of a renewal process model for chiasmata. The renewal process class of models includes almost all of the map functions proposed in the literature, but not the Sturt map function.

(3) Although the Sturt count-location model has no renewal process analogue, the *truncated Poisson* distribution does (Browning, 1999). This shows that two quite different processes can lead to same map function (Speed, 1996). Further, Browning (1999) has shown that a zero-modified Poisson distribution is the *unique* model that is both a count-location and a stationary renewal-process. (This includes, of course, both the Poisson model and the truncated Poisson model.)

(4) A flexible and simple renewal-process model is the *chi-square model* (Zhao et al., 1995). The renewal density is a scaled $\chi^2_{2(m+1)}$, with the scaling $(4(m+1))^{-1}$ such that the expected inter-arrival distance is $\frac{1}{2}$. One interpretation of this model is that potential chiasmata occur as a Poisson process and that every $(m+1)^{\text{th}}$ such potential chiasma becomes an actual chiasma. These models fit data well

(Zhao et al., 1995), and have properties that make recombination probabilities over several loci, and hence likelihood computations on pedigrees, somewhat tractable (Lin and Speed, 1996). A generalization of the chi-squared model is the *Poisson-skip model* (Lange, 1997). In this case, the r th potential chiasmata becomes one with probability β_r. The renewal density is a mixture of chi-squared (χ^2) distributions, with the scaling of genetic distance again chosen such the mean inter-arrival time of the chiasma process on the tetrad is $\frac{1}{2}$.