

Problem Set 1  
Solutions

Special topics in genetics and genomics (140.668)

1. This is rather a cute problem. If you toss  $n \geq 1$  fair coins, the chance of obtaining an odd number of heads is  $1/2$ .

The simplest solution is to use induction.

The cases  $n = 0$  and  $n = 1$  are obvious.

Suppose you toss  $n + 1$  fair coins, where  $n \geq 1$ .  $\Pr(\text{odd number of heads in } n + 1 \text{ tosses}) = \Pr(\text{odd number of heads in } n \text{ tosses}) \times \Pr(n + 1 \text{st toss is tails}) + \Pr(\text{even number of heads in } n \text{ tosses}) \times \Pr(n + 1 \text{st toss is heads}) = \dots = 1/2$ .

2. Consider an interval of length  $d$  on a chromosome of length  $L$ .

Let  $A = \{\text{no chiasma in interval}\}$  and  $N = \text{total no. chiasmata on the chromosome}$ .

$$\begin{aligned} \text{Map function: } M(d) &= [1 - \Pr(A)]/2 \\ &= [1 - \sum_{n=0}^{\infty} \Pr(N = n \text{ and } A)]/2 \\ &= [1 - \sum_{n=0}^{\infty} \Pr(N = n) \Pr(A|N = n)]/2 \\ &= [1 - \sum_{n=0}^{\infty} p_n (1 - d/L)^n]/2 \end{aligned}$$

In the case  $p_n = e^{-2L} (2L)^n / n!$ , we obtain:

$$\begin{aligned} M(d) &= [1 - \sum_{n=0}^{\infty} \frac{e^{-2L} (2L)^n}{n!} (1 - d/L)^n]/2 \\ &= [1 - \sum_{n=0}^{\infty} \frac{1}{n!} e^{-2L} [2L(\frac{L-d}{L})]^n]/2 \\ &= [1 - e^{-2d} \sum_{n=0}^{\infty} \frac{1}{n!} e^{-2(L-d)} [2(L-d)]^n]/2 \\ &= [1 - e^{-2d}]/2 \end{aligned}$$

3. Let the chromosome be represented by the interval  $[0, L]$ .

Consider any finite set of disjoint subintervals  $I_1, I_2, \dots, I_k$ . Let  $n_i = \text{no. chiasmata in interval } I_i$  (on the four-strand bundle) and  $m_i = \text{no. crossovers in } I_i$  (on a random meiotic product).

We wish to show that the  $\{m_i\}$  are independent and  $m_i \sim \text{Poisson}(|I_i|)$ .

Since the chiasma process is a Poisson process, the  $\{n_i\}$  are independent with  $n_i \sim \text{Poisson}(2|I_i|)$ .

Under no chromatid interference (NCI), the chiasmata are “thinned” independently with probability  $1/2$  to get the crossover process. Since the  $\{n_i\}$  are independent and since the thinning in the disjoint subintervals are independent, it should be clear that the  $\{m_i\}$  are independent.

So, we have  $m_i|n_i \sim \text{Binomial}(n_i, 1/2)$  and  $n_i \sim \text{Poisson}(2|I_i|)$ , and we need to show  $m_i \sim \text{Poisson}(|I_i|)$ . Let  $d = |I_i|$  and drop the subscripts  $i$ , to save a few keystrokes.

$$\begin{aligned}
\Pr(m = j) &= \sum_{k=j}^{\infty} \Pr(n = k \text{ and } m = j) \\
&= \sum_{k=j}^{\infty} \Pr(n = k) \Pr(m = j | n = k) \\
&= \sum_{k=j}^{\infty} \frac{e^{-2d} (2d)^k}{k!} \binom{k}{j} \left(\frac{1}{2}\right)^k \\
&= \sum_{k=j}^{\infty} \frac{e^{-2d} d^k}{k!} \left( \frac{k!}{j!(k-j)!} \right) \\
&= \left( \frac{e^{-d} d^j}{j!} \right) \sum_{k=j}^{\infty} \frac{e^{-d} d^{k-j}}{(k-j)!} \\
&= e^{-2d} d^j / j!
\end{aligned}$$

So  $m_i \sim \text{Poisson}(|I_i|)$ , and we're done.

4. Assume that the first chiasma involves a random pair of non-sister chromatids, that the next chiasma involves exactly the opposite pair, and subsequent chiasmata alternate in their choice of strands. This is strong chromatid interference.

Assume that the locations of chiasmata on the four-strand bundle are according to a stationary renewal process with inter-arrival distribution  $\text{gamma}(\text{shape}=1/2, \text{rate}=1)$ . This is strong negative chiasma interference.

Recall that if  $X_1, X_2 \sim \text{iid } \text{gamma}(\text{shape}=1/2, \text{rate}=1)$ , then  $X_1 + X_2 \sim \text{gamma}(\text{shape}=1, \text{rate}=1) \equiv \text{exponential}(\text{rate}=1)$ .

It should be clear, then, that the locations of crossovers on a random meiotic product follow a Poisson process, and so exhibit no crossover interference.

*What do we conclude?* We've seen that different combinations of chromatid interference and chiasma interference can give precisely the same crossover process on a random meiotic product. Thus, "single-spore" data (that is, data on random products of meiosis) are not sufficient to infer what's going on with both strand choice and chiasma location.