QTL mapping: other single-QTL models

1. Consider mapping a simple Mendelian disease gene using a backcross. Suppose you have essentially complete genotype data, so that the exact locations of crossovers can be determined. Let L denote the length (in cM) of the interval to which you will have mapped the gene with n backcross individuals (i.e., the distance between the two closest flanking recombination events).

   (a) How many backcross individuals will you need so that $\mathrm{E}(L) \le 1$ cM?

   (b) How many backcross individuals will you need so that $\Pr(L \le 1 \text{ cM}) \ge 80\%$?

2. Derive the extension of the Kruskal-Wallis statistic discussed in lecture, for nonparametric interval mapping.

   Let R_i denote the ranks of the phenotypes, and let p_{ig} denote the probability that mouse i has QTL genotype g, given the available marker data. Define $\bar{S}_g = \sum_i R_i p_{ig} / \sum_i p_{ig}$

   (a) Show that $\mathrm{E}(\bar{S}_g \mid \text{marker data, no linkage}) = (n+1)/2$.

   (b) Show that

   $$\mathrm{var}(\bar{S}_g \mid \text{marker data, no linkage}) = \left[\frac{n+1}{12}\right] \left[\frac{n\sum_i p_{ig}^2 - (\sum_i p_{ig})^2}{(\sum_i p_{ig})^2}\right]$$

   (c) Put it all together.