BIOSTAT551 Statistical Genetics II: Quantitative Traits **Winter 2004**

Homework 2

General Instructions: For all problems, show how you determined your answer. Be sure to define all variables that you use and justify your logic. If you have used software to obtain answers, output should be attached, but this is not a substitute for showing the steps of your analysis! The required data sets are available from the course web-site.

2.1) A study of trait values found that, for the trait of interest, the additive genetic variance was 72, and the additive-additive genetic variance was 5.3.

a) Estimate the genetic covariance between the trait values for a pair of half-siblings in this population.

b) Estimate the genetic covariance between the trait values for a pair of second cousins from this population.

c) Consider estimating the genetic covariance between the trait values for a pair of full-sibs from this population. Is this possible with only the information given above? What other information would you need?

2.2) As part of a study investigating the genetics of height in Cairn Terriers (this is the breed of Toto in the Wizard of Oz), twenty sib-pairs (males) were measured.

a) Assuming that the additive genetic variance is the dominant source of phenotypic covariance for height in Cairn Terriers, estimate the (narrow-sense) heritability of this trait.

b) Now assume that the first ten sib-pairs (pairs 1-10) consisted of dogs that been raised on a standard diet of dry dog food while the remaining pairs (11-20) were raised on "Dr. Bob's Miracle Grow Munchies". Use linear regression (or another technique) to estimate the effect of the "Miracle Grow" dog food, then subtract out that effect and redo the analysis you did in part A. What has happened to your estimate of the heritibility? Why?

2.3) The data set for this problem consists of pest-resistance scores for 100 corn plants. Perform a likelihood ratio test to determine whether there appears to be a major QTL associated with this trait. Specifically, you should be comparing a no-QTL model (i.e. all trait values come from a single normal distribution) with a one-QTL model (i.e. the trait values come from a mixture of three normal distributions). You can use the NOCOM program for this exercise or write your own program. Note that the program will need to use an EM algorithm to get the maximum likelihood estimators for gene frequencies, means of the three normal distributions, and variance of the three normal distributions. This is explained in detail in Appendix 4 of the text. You may assume that the population is in Hardy-Weinburg equilibrium and that (in the case of the one-QTL model) all three normal distributions have the same variance. When you have done the analysis, answer the following questions:

a) For the one-QTL model, what are the maximum-likelihood estimates for the proportions of individuals in each genotypic class (QQ, Qq, and qq)? What are the mean scores for corn plants in each class? Based on your findings, does the trait appear to be strictly additive?

b) What is the value of your chi-square test statistic? How many degrees of freedom?

c) What is the p-value for your test? What do you conclude?

2.4) This problem uses the data from example 4.3 of handout 4. Using the summarized results, what do you conclude regarding heritability of cellular radiosensitivity? Your answer should include a discussion of the various models, including what evidence can be supported STATISTICALLY. In addition, comment on the presence of major genetic effects, polygenic background, environmental components and Mendelian transmission. Finally, conclude your discussion by providing the most parsimonious model and stating what you would consider the next logical step for a follow-up study.

2.5) Project Outlines. We are off to a good start with a number of interesting project proposals. Now it is time to lay the foundation for what will become your project paper. Please provide a **detailed outline** for your project that includes the information below. A bulleted outline is fine as long it includes each of the items below. You may also want to see the project description link from the course web-site to see the grading rubric that will be used for the final project.

Introduction: Summarize the project topic and include a statement of the importance of the topic as well as a statement of its relevance to statistical methods in quantitative genetics

Background and literature review: This should summarize the materials, papers, books etc. that you will use for your background and literature review. At this time a citation with a brief summary of the contained material will suffice. This should be a firm list of resources i.e. make sure you have pulled together the resources that you will need to finish the project in full.

Description of statistical methods: This should summarize the materials, papers, books etc. for the methods used in your project. At this time a citation with a brief summary of the contained material will suffice. Again, this should be a firm list of resources i.e. make sure you have pulled together the resources that you will need to finish the project in full.

Results: This should include a description of the major results that are relevant for your project. Do not include the results themselves only a summary description for each result!

Discussion: If possible, include a description of how this particular topic/paper improves upon and is different from previous work. If you are not comfortable with this yet, provide your current thoughts as to why this is an improvement. Be sure to leave bullet points for 1) additional research suggested by the authors and 2) additional research suggested by you. Provide the latter information if you have developed the project to that point.