Biometry 755 HW #4 KEY

The questions in this homework pertain to Table 2 on page 70 of the article:
Kishi et al. Blood 2004; 103:67 – 72
"Effects of prednisone and genetic polymorphisms on etoposide disposition in children with acute lymphoblastic leukemia"

We will use the following model:

$$\log(\text{etoposide clearance}) = \beta_0 + \beta_1 \text{Age} + \beta_2 R + \beta_3 M_1 + \beta_4 M_2 + \beta_5 C_1 + \beta_6 C_2 + \beta_7 G_1 + \beta_8 G_2 + \beta_9(\text{Age} \times R) + \beta_{10}(C_1 \times R) + \beta_{11}(C_2 \times R) + \beta_{12}(G_1 \times R) + \beta_{13}(G_2 \times R) + \varepsilon$$

where

RACE indicator: R = 1 if black
R = 0 if white

MDR1 indicators: M_1 = 1 if TT, M_1 = 0 if otherwise; M_2 = 1 if CT, M_2 = 0 if otherwise

CYP3A5 indicators: C_1 = 1 if AG, C_1 = 0 if otherwise; C_2 = 1 if GG, C_2 = 0 if otherwise

GSTP indicators: G_1 = 1 if AG, G_1 = 0 if otherwise; G_2 = 1 if GG, G_2 = 0 if otherwise

To create race specific models from this one, substitute R=0 or R=1 as appropriate.

1. *Consider the coefficients for the explanatory variable GSTP1*

   a. *The effect of the GSTP1 polymorphism on log etoposide clearance is measured relative to which genotype?*

      AA genotype -- from Table 2, we see that AA is listed as the baseline (reference) category for this set of indicator variables.

   b. *What is the estimated effect of the AG genotype on log etoposide clearance relative to the reference genotype among black subjects? This effect is an estimate of what model parameter(s) stated in question 1? Interpret this model parameter (on the scale of the log etoposide clearance).*

      For black subjects, the estimated effect of the AG genotype on log etoposide clearance relative to the reference genotype AA is 0.07.

      The parameters involved in modeling the AG genotype among black subjects are β_7 (since G_1 = 1 if AG) and β_{12} (since both R = 1 and G_1 = 1 for black subjects with AG). Therefore, 0.07 is an estimate for $\beta_7 + \beta_{12}$.

      Among black subjects, there is a 0.07 increase in the log etoposide clearance for individuals with the AG genotype compared to the AA genotype, where the rate of etoposide clearance is measured by ml/min/m^2.

c. *What is the estimated effect of the GG genotype on log etoposide clearance relative to the reference genotype among black subjects? This effect is an estimate of what model parameter(s) stated in question 1? Interpret this model parameter (on the scale of the log etoposide clearance).*

Among black subjects, the estimated effect of the GG genotype on log etoposide clearance relative to the reference genotype AA is 0.22

The parameters involved in modeling the GG genotype among black subjects are β_8 (since $G_2 = 1$ if GG) and β_{13} (since both R = 1 and $G_2 = 1$ for black subjects with GG). Therefore, 0.22 is an estimate for $\beta_8 + \beta_{13}$.

Among black subjects, there is a 0.22 increase in the log etoposide clearance comparing GG genotype to AA genotype, where the rate of etoposide clearance is measured by ml/min/m^2.

d. *What is the estimated effect of the AG genotype on log etoposide clearance relative to the reference genotype among white subjects? This effect is an estimate of what model parameter(s) stated in question 1? Interpret this model parameter (on the scale of the log etoposide clearance).*

Among white subjects, the estimated effect of the AG genotype on log etoposide clearance relative to the reference genotype AA is – 0 .06.

The parameters involved in modeling the AG genotype among white subjects are β_7 (since $G_1 = 1$ if AG). For whites, we do not pick up an additional race interaction term, since white corresponds to R = 0. Therefore, -0.06 is an estimate for β_7.

Among white subjects, there is a 0.06 decrease in the log etoposide clearance comparing AG genotype to AA genotype, where the rate of etoposide clearance is measured by ml/min/m^2.

e. *What is the estimated effect of the GG genotype on log etoposide clearance relative to the reference genotype among white subjects? This effect is an estimate of what model parameter(s) stated in question 1? Interpret this model parameter (on the scale of the log etoposide clearance).*

For white subjects, the estimated effect of the GG genotype on log etoposide clearance relative to the reference genotype AA is – 0.07.

The parameters involved in modeling the GG genotype among white subjects are β_8 (since $G_2 = 1$ if GG). Therefore, -0.07 is an estimate for β_8.

Among white subjects, there is a 0.07 decrease in the log etoposide clearance comparing GG genotype to AA genotype, where the rate of etoposide clearance is measured by ml/min/m^2.

2. *The footnote of Table 2 states "P values for terms interacting with race were obtained by F test using race-specific contrasts, after adjusting for other predictors in the model." Let's determine where the authors obtained the p-values for tests associated with white subjects.*

a. *Based on the model written in Question 1 of the class discussion, write out the model for log etoposide clearance for white subjects. Be sure to simplify expressions by combining appropriate like terms. (You can just use the one we wrote in class.)*

For white subjects: R = 0

$$\log(\text{etoposide clearance}) = \beta_0 + \beta_1 \text{Age} + \beta_2(0) + \beta_3 M_1 + \beta_4 M_2 + \beta_5 C_1 + \beta_6 C_2 + \beta_7 G_1 + \beta_8 G_2 + \beta_9(\text{Age} \times 0) + \beta_{10}(C_1 \times 0) + \beta_{11}(C_2 \times 0) + \beta_{12}(G_1 \times 0) + \beta_{13}(G_2 \times 0) + \varepsilon$$

$$\log(\text{etoposide clearance}) = \beta_0 + \beta_1 * \text{age} + \beta_3 * M_1 + \beta_4 * M_2 + \beta_5 * C_1 + \beta_6 * C_2 + \beta_7 * G_1 + \beta_8 * G_2 + \varepsilon$$

b. *Based on the model for white subjects, state an appropriate null and alternative hypothesis to test the association between age and log etoposide clearance among white subjects.*

H_0: $\beta_1 = 0$ (all other parameters $\neq 0$)
H_1: $\beta_1 \neq 0$ (all other parameters $\neq 0$)

c. *What test have we learned about in class that could be used to test the significance of association between age and log etoposide clearance among white subjects? What is the distribution of the test statistic for this test under the null hypothesis? What p-value stated in Table 2 corresponds to this test?*

Partial F-test ~ $F_{1, 88}$
The numerator df for this F-test comes from the number of contrasts we are testing (ie: how many parameters we are setting to 0 under the null hypothesis). In this case, we are only testing one parameter, and thus our numerator df = 1. The denominator df is given by n - (k+1), where (k+1) represents the number of parameters in the model. We see from the footnote of Table 2 that n = 102, and our full model has 14 parameters. Thus, the denominator df = 102 - 14 = 88.

The p- value associated with age for white subjects is 0.275.

d. *Based on the model for white subjects, state an appropriate null and alternative hypothesis to test the association between CYP3A5 polymorphisms and log etoposide clearance among white subjects.*

H_0: $\beta_5 = \beta_6 = 0$ (all other parameters $\neq 0$)
H_1: at least one of β_5 and $\beta_6 \neq 0$ (all other parameters $\neq 0$)

e. *What test have we learned about in class that could be used to test the significance of association between CYP3A5 polymorphisms and log etoposide clearance among white subjects? What is the distribution of the test statistic for this test under the null hypothesis? What p-value stated in Table 2 corresponds to this test?*

Multiple partial F-test ~ $F_{2, 88}$
Now, we are testing two parameters, so our numerator df = 2. Again, the denominator df = 102 - 14 = 88.

The p-value associated with CYP3A5 for white subjects is 0.862.

f. *Based on the model for white subjects, state an appropriate null and alternative hypothesis to test the association between GSTP1 polymorphisms and log etoposide clearance among white subjects.*

H_0: $\beta_7 = \beta_8 = 0$ (all other parameters $\neq 0$)
$H_{1:}$ at least one of β_7 and $\beta_8 \neq 0$ (all other parameters $\neq 0$)

g. *What test have we learned about in class that could be used to test the significance of association between GSTP1 polymorphisms and log etoposide clearance among white subjects? What is the distribution of the test statistic for this test under the null hypothesis? What p-value stated in Table 2 corresponds to this test?*

Multiple partial F-test ~ $F_{2, 88}$
Again, we are testing two parameters, so our numerator df = 2. The denominator df = 102 - 14 = 88.

The p- value associated with GSTP1 for white subjects is 0.198.

3. *Interpret the coefficients in Table 2 corresponding to genetic polymorphisms of GSTP1 among black subjects on the original scale of measurement, i.e. on the scale of etoposide clearance. Demonstrate (or explain) how you obtain this interpretation.*

The coefficients are derived from the model of the log(etoposide clearance). In order to transform these coefficients back to our original scale of measurement (etoposide clearance rate), we will take the antilog.

AG:
log coefficient = 0.07
original scale coefficient = $e^{0.07} = 1.07$
Among black subjects, there is a 7% increase in the rate of etoposide clearance comparing AG genotype to AA genotype, where the rate of etoposide clearance is measured by ml/min/m^2.

GG:
log coefficient = 0.22
original scale coefficient = $e^{0.22} = 1.25$
Among black subjects, there is a 25% increase in the rate of etoposide clearance comparing GG genotype to AA genotype, where the rate of etoposide clearance is measured in ml/min/m^2.