## HOMEWORK 6 — SOLUTIONS

**Problem 7.2**

(a) We have

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix}, X^TX = \begin{pmatrix} n & 0 & n \\ 0 & n & 0 \\ n & 0 & Cn \end{pmatrix}, (X^TX)^{-1} = \frac{1}{n} \begin{pmatrix} \frac{C}{C-1} & 0 & -\frac{1}{C-1} \\ 0 & 1 & 0 \\ -\frac{1}{C-1} & 0 & \frac{1}{C-1} \end{pmatrix},$$

so the normal equations lead to

$$\begin{aligned} n\hat\beta_0 &= \frac{C}{C-1}\sum y_i - \frac{1}{C-1}\sum y_i x_i^2, \\ n\hat\beta_1 &= \sum y_i x_i, \\ n\hat\beta_2 &= -\frac{1}{C-1}\sum y_i + \frac{1}{C-1}\sum y_i x_i^2. \end{aligned}$$

The covariance matrix is $(X^TX)^{-1}\sigma^2$.

(b) Writing $g(\beta_0, \beta_1, \beta_2) = -\frac{\beta_1}{2\beta_2}$, we have $\frac{\partial g}{\partial \beta_0} = 0$, $\frac{\partial g}{\partial \beta_1} = -\frac{1}{2\beta_2}$, $\frac{\partial g}{\partial \beta_2} = \frac{\beta_1}{2\beta_2^2}$. The variance of $\hat\theta$ is approximately $\frac{\sigma^2}{n}\left\{\left(\frac{\partial g}{\partial \beta_1}\right)^2 + \frac{1}{C-1}\left(\frac{\partial g}{\partial \beta_2}\right)^2\right\}$, which reduces to $\frac{\sigma^2}{4n\beta_2^2}\left\{1 + \frac{\beta_1^2}{(C-1)\beta_2^2}\right\}$.

(c) The confidence interval is $\hat\theta \pm \frac{t^* s}{2\sqrt{n}\beta_2}\sqrt{1 + \frac{\beta_1^2}{(C-1)\beta_2^2}}$ where $t^*$ is the $(1-\alpha/2)$ point of the $t_{n-3}$ distribution.

(d) Rewriting the response as $\gamma_0 + \gamma_1(x - x^2) + \gamma 2x$, this has derivative $\gamma_1(1-2x) + \gamma_2$, which is 0 at $x = \frac{1}{2}$ if and only if $\gamma_2 = 0$ — this therefore becomes the hypothesis we would like to test.
Under the new parameterization, the $X$, $X^TX$ and $(X^TX)^{-1}$ matrices become

$$X = \begin{pmatrix} 1 & x_1 - x_1^2 & x_1 \\ 1 & x_2 - x_2^2 & x_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n - x_n^2 & x_n \end{pmatrix}, X^TX = n\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1+C & 1 \\ 0 & 1 & 1 \end{pmatrix}, (X^TX)^{-1} = \frac{1}{Cn}\begin{pmatrix} C & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & C \end{pmatrix},$$

where the calculation of $(X^TX)^{-1}$ used the hint.
Thus, the estimate of $\hat\gamma_2$ is $\frac{1}{C-1}\sum y_i\{x_i^2 + (C-1)x_i - 1\}$ and its standard error is $s\sqrt{\frac{C}{C-1}}$. The test of level $\alpha$ will reject $H_0$ whenever $|\hat\gamma_2|/\left(s\sqrt{\frac{C}{C-1}}\right) > t_{n-3,1-\alpha/2}$.

(e) The hypothesis is $-\frac{\beta_1}{\beta_2} = -\frac{\delta_1}{\delta_2}$ or equivalently $\theta = \beta_1\delta_2 - \beta_2\delta_1 = 0$. This is not a linear hypothesis to it is not possible to find an exact test. We can estimate $\hat{\theta} = \hat{\beta}_1\hat{\delta}_2 - \hat{\beta}_2\hat{\delta}_1$ with standard error approximately $\sqrt{\left(\frac{\partial\theta}{\partial\beta_1}\right)^2 \mathrm{Var}(\hat{\beta}_1) + \left(\frac{\partial\theta}{\partial\beta_2}\right)^2 \mathrm{Var}(\hat{\beta}_2) + \left(\frac{\partial\theta}{\partial\delta_1}\right)^2 \mathrm{Var}(\hat{\delta}_1) + \left(\frac{\partial\theta}{\partial\delta_2}\right)^2 \mathrm{Var}(\hat{\delta}_2)} \approx \frac{s}{\sqrt{n}}\sqrt{(\hat{\beta}_1^2 + \hat{\delta}_1^2)/(C-1) + \hat{\beta}_2^2 + \hat{\delta}_2^2}$. Here we are assuming the error variance $\sigma^2$ is the same for both halves of the experiment and is estimated by $s^2$. The test will be based on the approximate $t_{n-3}$ distribution for $\hat{\theta}/SE(\hat{\theta})$.

**Problem 7.4**

(a) Prove by directly multiplying out so show $VW = (1-\rho^2)I$, so $\kappa = (1-\rho^2)^{-1}$.

(b) With $X = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \end{pmatrix}^T$, $Y = \begin{pmatrix} y_1 & y_2 & \ldots & y_n \end{pmatrix}^T$, we have (there are several equivalent notations)

$$X^T V^{-1} Y = \frac{1}{1-\rho^2}\left\{x_1 y_1 + x_n y_n + (1+\rho^2)\sum_{i=2}^{n-1} x_i y_i - \rho\sum_{i=1}^{n-1}(x_i y_{i+1} + x_{i+1} y_i)\right\}$$

and $X^T V^{-1} X$ the same on substituting $x_i$ for $y_i$. The GLS estimator is then $\hat{\beta} = X^T V^{-1} Y / X^T V^{-1} X$ and its variance is $\sigma^2 / X^T V^{-1} X$.

**Problem 7.5**

First calculate

$$X = \begin{pmatrix} 1 & -b & b^2 \\ 1 & -a & a^2 \\ 1 & 0 & 0 \\ 1 & a & a^2 \\ 1 & b & b^2 \end{pmatrix}, \quad X^T X = \begin{pmatrix} 5 & 0 & 2(a^2+b^2) \\ 0 & 2(a^2+b^2) & 0 \\ 2(a^2+b^2) & 0 & 2(a^4+b^4) \end{pmatrix},$$

$$(X^T X)^{-1} = \begin{pmatrix} \frac{2(a^4+b^4)}{\Delta} & 0 & \frac{-2(a^2+b^2)}{\Delta} \\ 0 & \frac{1}{2(a^2+b^2)} & 0 \\ \frac{-2(a^2+b^2)}{\Delta} & 0 & \frac{5}{\Delta} \end{pmatrix}$$

where $\Delta = 10(a^4+b^4) - 4(a^2+b^2)^2 = 2(3a^4 - 4a^2b^2 + 3b^4)$. Thus, in particular, $\mathrm{Var}(\hat{\beta}_1) = A\sigma^2$, $\mathrm{Var}(\hat{\beta}_2) = B\sigma^2$ and $\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2)$ is 0.

To test the hypothesis $H_0: \ \beta_1 + 2\beta_2 x = 0$, first form

$$t = \frac{\hat{\beta}_1 + 2\hat{\beta}_2 x}{s\sqrt{A + 4Bx^2}}.$$

We will accept $H_0$ at level $\alpha$ when $|t| < t^*$. This leads to the given test.

Rearranging terms, the critical values for $x$ solve the quadratic equation

$$4(\hat{\beta}_2^2 - t^{*2}s^2B)x^2 + 4\hat{\beta}_1\hat{\beta}_2 x + \hat{\beta}_1^2 - t^{*2}s^2A = 0$$

with roots

$$\frac{-\hat{\beta}_1\hat{\beta}_2 \pm \sqrt{(-\hat{\beta}_1\hat{\beta}_2)^2 - (\hat{\beta}_2^2 - t^{*2}s^2B)(\hat{\beta}_1^2 - t^{*2}s^2A)}}{2(\hat{\beta}_2^2 - t^{*2}s^2B)}$$

The roots are real if the argument of the square root sign is positive, and this occurs if $\hat{\beta}_1^2B + \hat{\beta}_2^2A - (t^*s)^2AB > 0$; in that case, the length of the confidence interval is

$$\frac{\sqrt{(-\hat{\beta}_1\hat{\beta}_2)^2 - (\hat{\beta}_2^2 - t^{*2}s^2B)(\hat{\beta}_1^2 - t^{*2}s^2A)}}{\hat{\beta}_2^2 - t^{*2}s^2B}$$

which quickly reduces to the form given.

**Problem 7.12**

The best model has $\hat{\theta}_1 = 260.7, \hat{\theta}_1 = 5.7 \times 10^5, \hat{\theta}_3 = 0.159$ with standard errors respectively $1.183, 1.4 \times 10^6, 0.037$. In this case the hypothesis $\theta_3 = 0$ is clearly rejected and it seems highly plausible that we could estimate $\theta_1$. Of course this does not allow for other reasons why the performances could have levelled off, including more rigorous drug testing and the fact that the mile race is no longer regularly competed in championship events.