Assignment Five

(1) $E(y) = E(X\beta+\varepsilon) = E(X\beta) + E(\varepsilon) = X\beta + 0 = X\beta$

$cov(y) = cov(X\beta+\varepsilon) = cov(\varepsilon) = \sigma^2 I_n$

(2) $\hat{\beta} = (X'X)^{-1}X'y$ is $(k+1)\times 1$

(3) $E(\hat{\beta}) = E((X'X)^{-1}X'y) = (X'X)^{-1}X'E(y)$

$= (X'X)^{-1}X'X\beta = \beta$. Yes, unbiased.

(4) $cov(\hat{\beta}) = cov((X'X)^{-1}X'y) = (X'X)^{-1}X' cov(y)\left((X'X)^{-1}X'\right)'$

$= (X'X)^{-1}X'\sigma^2 I_n X(X'X)^{-1} = \sigma^2(X'X)^{-1}X'X(X'X)^{-1}$

$= \sigma^2 (X'X)^{-1}$

(5) $\hat{y} = X\hat{\beta}$ $(n \text{ by } k+1)$ times $(k+1 \text{ by } 1) = n\times 1$

(6) $E(\hat{y}) = E(X\hat{\beta}) = XE(\hat{\beta}) = X\beta$

(7) $cov(\hat{y}) = cov(X\hat{\beta}) = X\,cov(\hat{\beta})X' = X\sigma^2(X'X)^{-1}X'$

$= \sigma^2 X(X'X)^{-1}X' = \sigma^2 H$

(8) $n\times 1$

(9) $E(\hat{\varepsilon}) = E(y-\hat{y}) = E(y) - E(\hat{y}) = X\beta - X\beta = 0$

(10) $cov(\hat{\varepsilon}) = cov((I-H)y) = (I-H)\,cov(y)\,(I-H)'$

$= (I-H)\sigma^2 I_n(I-H) = \sigma^2(I-H)(I-H)$

$= \sigma^2(I-H)$

2

11 (a) $E(\bar{Y}) = \mu$, $Var(\bar{Y}) = \frac{\sigma^2}{n}$

(b) $E(L) = \mu = E\left(\sum_{i=1}^{n} c_i Y_i\right) = \sum_{i=1}^{n} c_i E(Y_i)$

$$= \mu \sum_{i=1}^{n} c_i$$

Since L is unbiased, $\mu = \mu \sum_{i=1}^{n} c_i$ is true in particular for $\mu = 1$, so $\sum_{i=1}^{n} c_i = 1$.

(c) Yes; $c_i = \frac{1}{n}$ for $i = 1, \dots, n$.

(d) $Var(L) = Var\left(\sum_{i=1}^{n} c_i Y_i\right) \overset{ind}{=} \sum_{i=1}^{n} Var(c_i Y_i) = \sum_{i=1}^{n} c_i^2 \, Var(Y_i)$

$$= \sum_{i=1}^{n} c_i^2 \sigma^2 = \sigma^2 \sum_{i=1}^{n} c_i^2$$

(e) So to minimize $Var(L)$, minimize $\sum_{i=1}^{n} c_i^2$ subject to the constraint $\sum_{i=1}^{n} c_i = 1$.

$$\sum_{i=1}^{n} c_i^2 \overset{Hint}{=} \sum_{i=1}^{n} \left(c_i - \frac{1}{n} + \frac{1}{n}\right)^2 = \sum_{i=1}^{n} \left(\left(c_i - \frac{1}{n}\right)^2 + 2\left(c_i - \frac{1}{n}\right)\frac{1}{n} + n\left(\frac{1}{n^2}\right) \right)$$

$$= \sum_{i=1}^{n} \left(c_i - \frac{1}{n}\right)^2 + 2 \cdot \frac{1}{n} \sum_{i=1}^{n} \left(c_i - \frac{1}{n}\right) + \frac{1}{n}$$

$$= \sum_{i=1}^{n} \left(c_i - \frac{1}{n}\right)^2 + \frac{2}{n} \underbrace{\left(\sum_{i=1}^{n} c_i - 1\right)}_{\substack{=0 \\ \text{Bec } \sum c_i = 1}} + \frac{1}{n} = \sum_{i=1}^{n} \left(c_i - \frac{1}{n}\right)^2 + \frac{1}{n}$$

$\geq \frac{1}{n}$, with equality when $c_i = \frac{1}{n}$ for $i = 1, \dots, n$, in which case $L = \bar{Y}$.

(12) (a) $\ell'\hat{\beta}$

(b) $E(\ell'\hat{\beta}) = \ell' E(\hat{\beta}) = \ell'\beta$

(c) $\ell'\hat{\beta} = \ell'(X'X)^{-1}X'y = c_0'y$, so

$c_0 = X(X'X)^{-1}\ell$

(d) $\mathrm{Var}(c'y) = \mathrm{cov}(c'y) = c'\,\mathrm{cov}(y)\,c = c'\sigma^2 I_n c$

$= \sigma^2 c'c.$

(e) $\ell'\beta = E(c'y) = c'E(y) = c'X\beta = v'\beta$ for all $\beta \in \mathbb{R}^{k+1}$. So in particular, it's true for $\beta_{(1)} = (1,0\cdots0)'$, and $\ell'\beta_{(1)} = v'\beta_{(1)} \Rightarrow \ell_1 = v_1$

It's true for $\beta_{(2)} = (0,1,\cdots0)$, so $\ell_2 = v_2$

$\vdots$

It's true for $\beta_{(k+1)} = (0,0,\cdots,1)$, so $\ell_{k+1} = v_{k+1}$

And we have $\ell = v = (c'X)' = X'c$ qed

(f) $(c-c_0)'c_0 = (c - X(X'X)^{-1}\ell)'X(X'X)^{-1}\ell$

$= (c - X(X'X)^{-1}X'c)X(X'X)^{-1}X'c$

$= (c - Hc)'Hc = (c' - c'H')Hc$

$= c'Hc - c'H'Hc = c'Hc - c'HHc$

$= c'Hc - c'Hc = 0$

(12g) $c'c = (c - c_0 + c_0)'(c - c_0 + c_0)$

$$= (c-c_0)'(c-c_0) + \underbrace{(c-c_0)'c_0 + c_0'(c-c_0)}_{1\times 1,\ \text{and equal}} + c_0'c_0$$

$$= (c-c_0)'(c-c_0) + \underbrace{2(c-c_0)'c_0}_{=0 \text{ by part (f)}} + c_0'c_0$$

$$= (c-c_0)'(c-c_0) + c_0'c_0 \quad \text{as required}$$

(h) $(c-c_0)'(c-c_0) = z'z = \sum_{i=1}^{n} z_i^2 \geq 0$, and equals zero iff $z_i = 0$, for $i = 1, \ldots, n$ iff $z = c - c_0 = 0$

$\iff c = c_0$. So the minimum is unique.

(13) (a) $\mathrm{Var}(\hat{\beta}) = \mathrm{Var}\left(\dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}\right)$

$= \dfrac{1}{(\sum x_i^2)^2} \mathrm{Var}\left(\sum_{i=1}^{n} x_i y_i\right) \overset{\text{ind}}{=} \dfrac{1}{(\sum x_i^2)^2} \sum_{i=1}^{n} \mathrm{Var}(x_i y_i)$

$= \dfrac{1}{(\sum x_i^2)^2} \sum_{i=1}^{n} x_i^2 \mathrm{Var}(y_i) = \dfrac{1}{(\sum x_i^2)^2} \sum_{i=1}^{n} x_i^2 \sigma^2$

$= \dfrac{\sigma^2}{\sum_{i=1}^{n} x_i^2}$

(b) (i) $E(\hat{\beta}_2) = \dfrac{1}{\bar{x}} E(\bar{y}) = \dfrac{1}{\bar{x}} E\left(\dfrac{1}{n} \sum_{i=1}^{n} y_i\right)$

$= \dfrac{1}{\bar{x}} \dfrac{1}{n} \sum_{i=1}^{n} E(y_i) = \dfrac{1}{\bar{x}} \dfrac{1}{n} \sum_{i=1}^{n} E(\beta x_i + \varepsilon_i)$

$= \dfrac{1}{\bar{x}} \dfrac{1}{n} \sum_{i=1}^{n} (\beta x_i + 0) = \dfrac{1}{\bar{x}} \beta \dfrac{1}{n} \sum_{i=1}^{n} x_i = \dfrac{1}{\cancel{\bar{x}}} \beta \cancel{\bar{x}}$

$= \beta$ unbiased, as long as $\bar{x} \neq 0$

(ii) Yes. $\hat{\beta}_2 = \dfrac{1}{n\bar{x}} \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \left(\dfrac{1}{n\bar{x}}\right) y_i$, so

$c_i = \dfrac{1}{n\bar{x}}$ for $i = 1, \ldots, n$

(iii) $\mathrm{Var}(\hat{\beta}_2) \overset{\text{ind}}{=} \dfrac{1}{n^2 \bar{x}^2} \sum_{i=1}^{n} \mathrm{Var}(y_i) = \dfrac{n\sigma^2}{n^2 \bar{x}^2} = \dfrac{\sigma^2}{n\bar{x}^2}$

(iv) Gauss-Markov Theorem

(v) $\dfrac{\sigma^2}{\sum x_i^2} = \dfrac{\sigma^2}{n\bar{x}^2} \iff \sum_{i=1}^{n} x_i^2 = n\bar{x}^2 \iff$
$\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 = 0$, ALL x_i are equal

(13c) $\hat{\beta}_3 = \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{x_i}$

(i) $E(\hat{\beta}_3) = \frac{1}{n}\sum_{i=1}^{n} \frac{E(y_i)}{x_i} = \frac{1}{n}\sum_{i=1}^{n} \frac{\beta x_i}{x_i}$

$= \frac{n\beta}{n} = \beta$ unbiased as long as no $x_i = 0$

(ii) Yes: $c_i = \frac{1}{n x_i}$

(iii) $Var(\hat{\beta}_3) = \frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{x_i^2} Var(y_i)$

$= \frac{\sigma^2}{n^2}\sum_{i=1}^{n} \frac{1}{x_i^2}$

(iv) Gauss - Markov. $\frac{\cancel{\sigma}^2}{\sum_{i=1}^{n} x_i^2} \leq \frac{\cancel{\sigma}^2 \sum_{i=1}^{n} \frac{1}{x_i^2}}{n^2}$

$\iff \sum_{i=1}^{n} x_i^2 \geq \frac{n^2}{\sum_{i=1}^{n} \frac{1}{x_i^2}}$ A strange inequality, new to me, and very hard to prove otherwise.

(v) If all $x_i = 1$ (had to guess this) then $\hat{\beta} = \hat{\beta}_3 = \bar{y}$.

(14) (a) $Hv = HXb$, some $b \in \mathbb{R}^{k+1}$

$$= \underbrace{X(X'X)^{-1}X'X}_{I} b = Xb = v$$

(b) Every element of $X'\hat{\varepsilon} = 0$ is the inner product of a basis vector and $\hat{\varepsilon}$.

(c) $v'\hat{\varepsilon} = (Xb)'\hat{\varepsilon} = b'X'\hat{\varepsilon} = b'\underset{\sim}{0} = 0$

(d) The projection of $\hat{\varepsilon}$ onto $\mathcal{V}$ is the closest point

$$H\hat{\varepsilon} = H(y - \hat{y}) = Hy - H\hat{y} = Hy - HHy$$
$$= Hy - Hy = 0 \quad \underline{\text{Yes.}}$$

(e) $E(c'y) = \ell'\beta$ for all $\beta \in \mathbb{R}^{k+1}$ implies $\ell = X'c$, as in 12 (e). Then

$$Hc = X(X'X)^{-1}\underbrace{X'c}_{\ell} = X(X'X)^{-1}\ell = c_0$$

Projection is the closest point.

(15) (a) $y_i \sim N(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}, \sigma^2)$
$= N(x_i'\beta, \sigma^2)$

(b) The log likelihood is $\ell(\beta, \sigma^2)$

$$= \ln \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - x_i'\beta)^2}$$

$$= \ln\left(\sigma^{-n}(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2}\right)$$

$$= -n\ln\sigma - \frac{n}{2}\ln 2\pi \underbrace{- \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2}$$

To maximize this, maximize this part over β
Or equivalently, minimize

$$Q = \sum_{i=1}^{n}(y_i - x_i'\beta)^2 = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{1i} - \cdots - \beta_k x_{ik})^2$$

This is just the least squares problem, and it has the same solution.

(15c) We know that for every $\sigma^2 > 0$, the likelihood is maximized at $\beta = \hat{\beta}$. So maximize

$$\ell(\hat{\beta}, \sigma^2) = -\frac{n}{2}\ln\sigma^2 - \frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(y_i - x_i\vec{\beta})^2 (\sigma^2)^{-1}$$

$$= -\frac{n}{2}\ln\sigma^2 - \frac{n}{2}\ln(2\pi) - \frac{1}{2}\hat{\varepsilon}'\hat{\varepsilon}(\sigma^2)^{-1}$$ over σ^2.

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} - 0 - \frac{1}{2}\hat{\varepsilon}'\hat{\varepsilon}(-1)(\sigma^2)^{-2}$$

$$= -\frac{n}{2\sigma^2} + \frac{\hat{\varepsilon}'\hat{\varepsilon}}{2\sigma^4} \overset{set}{=} 0 \Rightarrow n = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2}$$

$\Rightarrow \sigma^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n}$. To check that it's a max,

$$\frac{\partial^2 \ell}{\partial \sigma^{2\,2}} = \frac{\partial}{\partial\sigma^2}\left(-\frac{n}{2}(\sigma^2)^{-1} + \frac{\hat{\varepsilon}'\hat{\varepsilon}}{2}(\sigma^2)^{-2}\right)$$

$$= -\frac{n}{2}(-1)(\sigma^2)^{-2} + \frac{\hat{\varepsilon}'\hat{\varepsilon}}{2}(-2)(\sigma^2)^{-3}$$

$$= \frac{n}{2\sigma^4} - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^6}$$ Evaluate at $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n}$

$$= \frac{n}{2\hat{\sigma}^4} - \frac{n\hat{\sigma}^2}{\hat{\sigma}^6} = \frac{n}{2\hat{\sigma}^4} - \frac{n}{\hat{\sigma}^4} < 0$$

concave down, max.

$s^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k-1}$, $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n}$, $\hat{\sigma}^2$ is biased but similar

```
> # Q16

>     #a) Calculate beta-hat
                  [,1]
Intercept 0.6080747411
VERBAL    0.0023070007
MATH      0.0009973607

>     #b) Predict GPA for Verbal=600, Math=700
[1] 2.690428

>     #c) Mean of yhat and mean of y
[1] 2.6301 2.6301

>     #d) epsilon-hat, vector of residuals
> mean(epsilonhat)
[1] -1.657453e-14

>     #e) Inner product of yhat and epsilonhat
              [,1]
[1,] -8.835738e-12

>     #f) Inner product of epsilonhat and total score is zero because total is in
the space spanned by the columns of X
              [,1]
[1,] -4.190341e-09
```

```
> # Q17: Faraway Ch. 2 Exercise 1, page 25.

>
> # (a) What percentage of variation in the response is explained by these
predictors?
[1] 0.5267234

> # (b) Which observation has the largest (positive) residual? Give the case
number.
      24
94.25222

> # (c) Compute the mean and median of the residuals.
[1] -3.065293e-17 -1.451392e+00

> # (d) Compute the correlation of the residuals with the fitted values.
[1] -1.070659e-16

> # (e) Compute the correlation of the residuals with the income.
[1] -7.242382e-17

> # (f) For all other predictors held constant, what would be the difference in
predicted expenditure on gambling for a male compared to a female?
      sex
-22.11833
> # 22 pounds LESS for female.
```