Problem 4 is due at class on Wednesday 28th October.

1. For a simple linear regression model with

show that

(a)  $\sum_{i=1}^{n} \widehat{\epsilon_i} = 0$ (b)  $\sum_{i=1}^{n} \widehat{\epsilon_i} X_i = 0$ 

2. Show that the expected value of MS(Reg) for the standard SLR model is

$$\sigma^2 + \beta_1^2 S_{XX}$$
.

3. A linear regression model may be written either

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \qquad i = 1, \dots, n$$

$$Y_i = \alpha_0 + \alpha_1(X_i - \bar{X}) + \epsilon_i$$
  $i = 1, \dots, n$ .

- (a) Find the relationship between the  $\alpha$ 's and the  $\beta$ 's.
- (b) Use the method of least squares to estimate  $\alpha_0$  and  $\alpha_1$ .
- 4. A random variable Y is distributed with expectation depending linearly on another variable X and with constant variance  $\sigma^2$ . Observations are denoted by  $(X_i, Y_i), i = 1, \dots, n$ . If we write

$$\mathrm{E}\{Y_i\} = \alpha_0 + \alpha_1(X_i - \bar{X})$$

- (a) Find the expected values and variances of the least squares estimators  $\widehat{\alpha}_0$  and  $\widehat{\alpha}_1$ .

$$\operatorname{Cov}\{Y_i, \widehat{\alpha}_1\} = \frac{\sigma^2(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$$

- (c) Hence or otherwise prove that  $\hat{\alpha}_0$  and  $\hat{\alpha}_1$  are uncorrelated and interpret this property.
- 5. A random variable Y has mean  $\alpha + \beta X$  and variance  $\sigma^2$ . At each of n values of X, a value of Y is observed giving n pairs of independent observations  $(X_1,Y_1),\ldots,(X_n,Y_n)$ . Suppose that the observations are made on two different days,  $n_1$  on the first day and  $n-n_1$  on the second. Suppose that there are practical reasons for assuming that conditions vary from day to day in such a way as to affect  $\alpha$  but not  $\beta$  or  $\sigma^2$ . Show that the least squares estimator of the common  $\beta$  derived from the two samples is given by

$$\widehat{\beta} = \frac{\omega_1 \widehat{\beta}_1 + \omega_2 \widehat{\beta}_2}{\omega_1 + \omega_2}$$

where  $\hat{\beta}_1$  and  $\hat{\beta}_2$  are the least squares estimators of  $\beta$  from the samples on the first and second days respectively, and  $\omega_1, \omega_2$  are the corresponding sums of squares of the X's about their means. Find  $\text{Var}\{\hat{\beta}\}$ .

6. Given a random sample of n data pairs  $(X_1, Y_1), \dots, (X_n, Y_n)$  a data analyst decides to draw a line on the scatterplot by joining the first and last points. Write out a formula for the slope of this line which will be scatterplot by joining the first and last points. Write out a formula for the slope of this line which will be scatterplot by joining the first and last points.

$$E(Y) = \alpha + \beta X$$

is  $\vec{\beta}$  an unbiased estimator of  $\beta$ ? How does the variance of  $\vec{\beta}$  compare with the variance of the least squares estimator  $\vec{\beta}$ ?

7. (\*) Consider the simple linear regression model

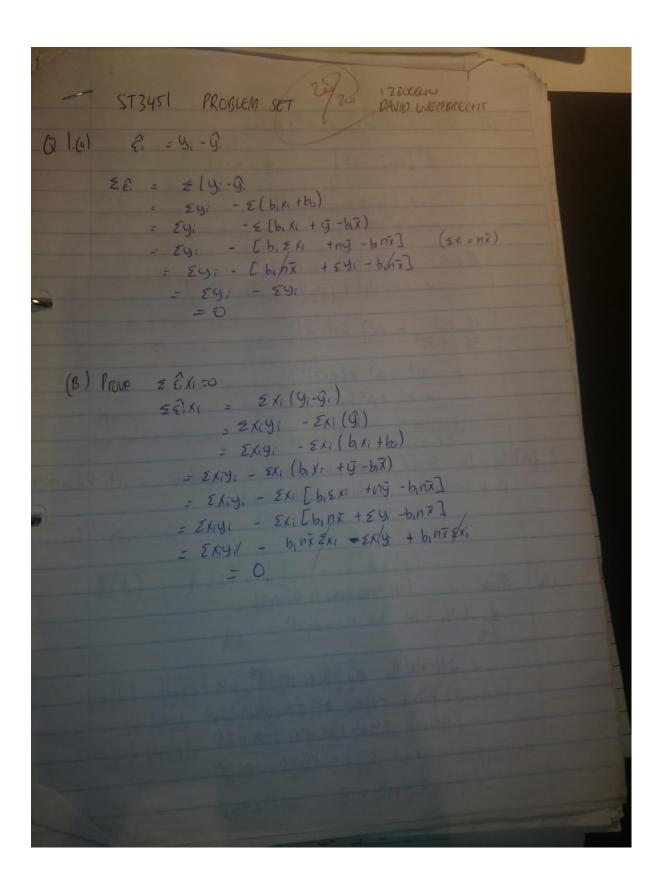
$$Y_i = \theta X_i + \epsilon_i$$
  $i = 1, \dots, n$ 

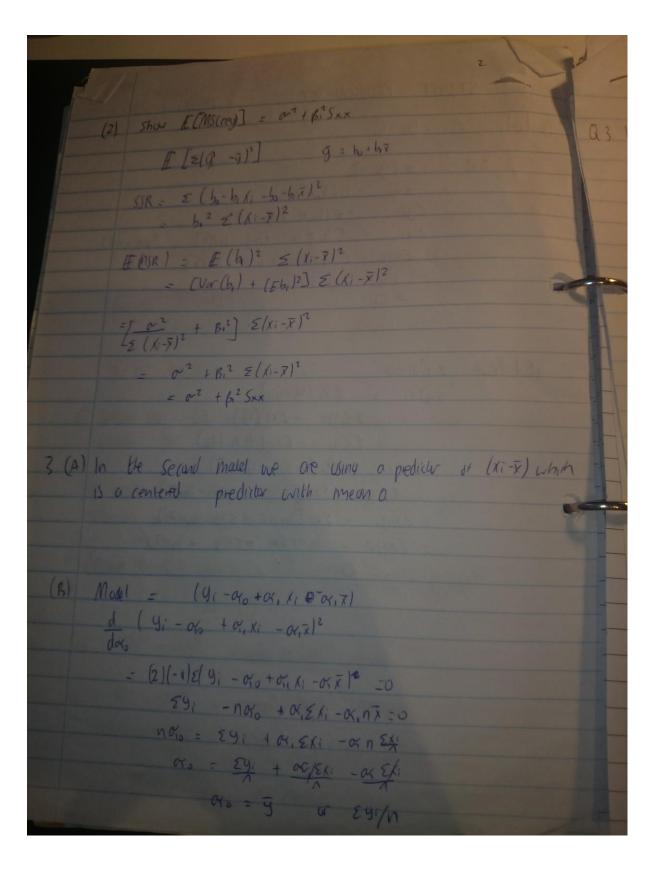
where the  $X_t$  are fixed constants and the  $\epsilon_t$  are uncorrelated normal variables with zero means and the same variance  $\sigma^2$ . Show how a confidence interval for  $\theta$  may be obtained based on the Student-t variable

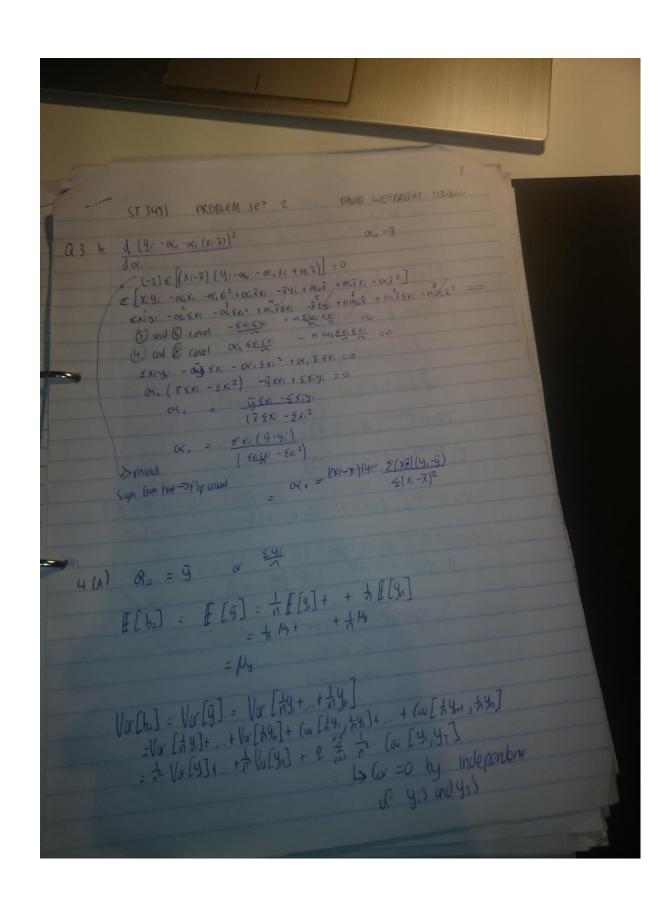
$$t_1 = \frac{\widehat{\theta} - \theta}{\left(s^2 / \sum X_i^2\right)^{1/2}}$$

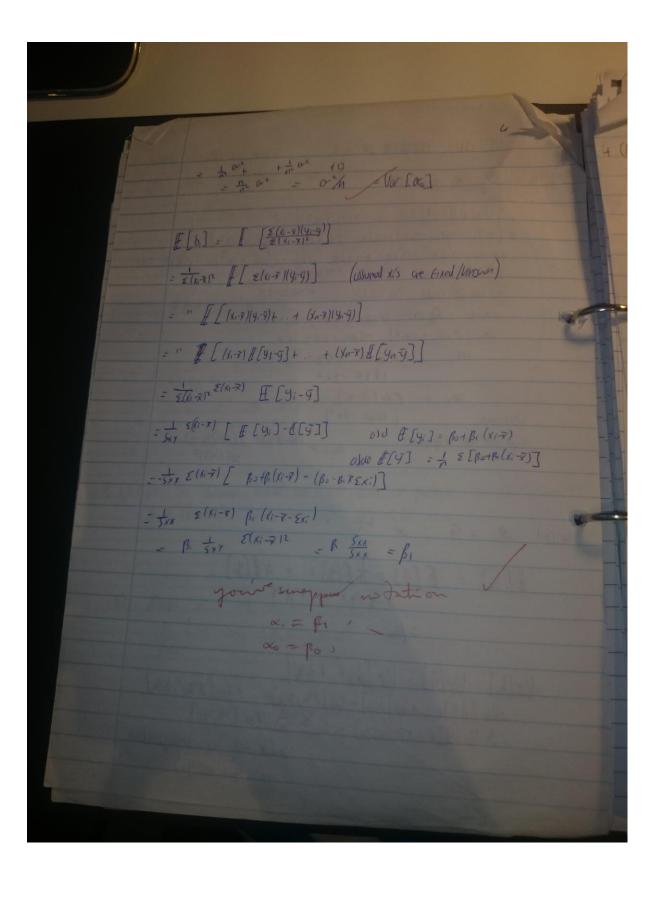
where  $\hat{\theta}$  is the least squares unbiased estimator of  $\theta$  and  $s^2$  is the usual unbiased estimator of  $\sigma^2$ . Show that this confidence interval is always of smaller length than the alternative interval based on the variable

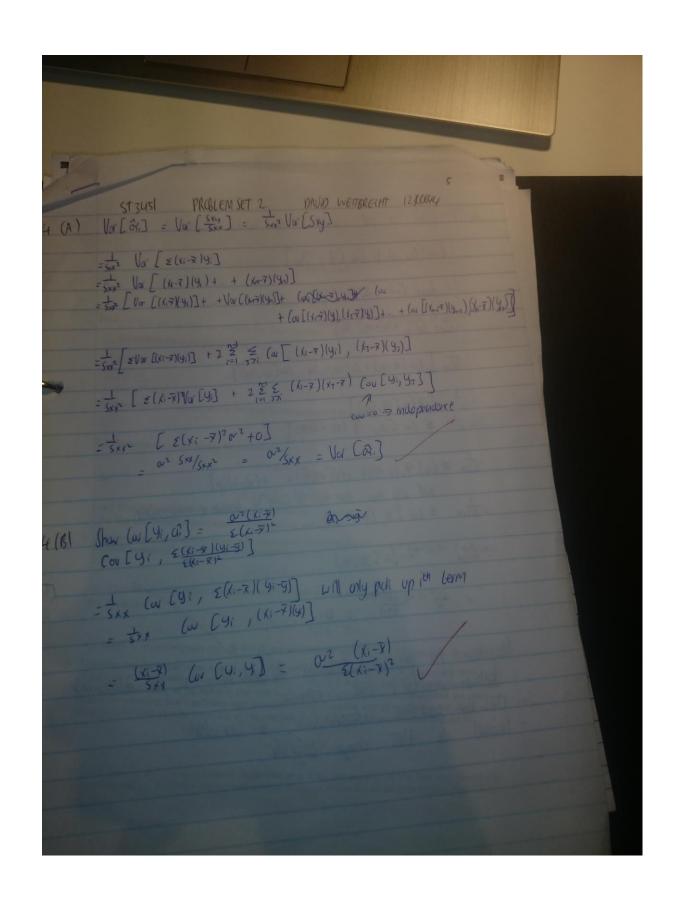
$$t_2 = \frac{\tilde{Y} - \theta \bar{X}}{(s^2/n)^{1/2}}$$



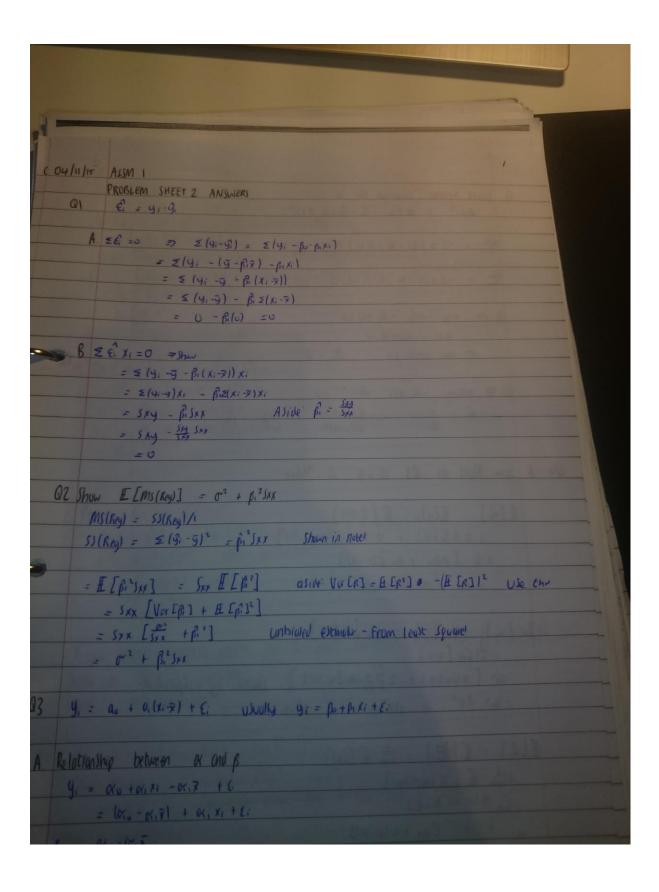




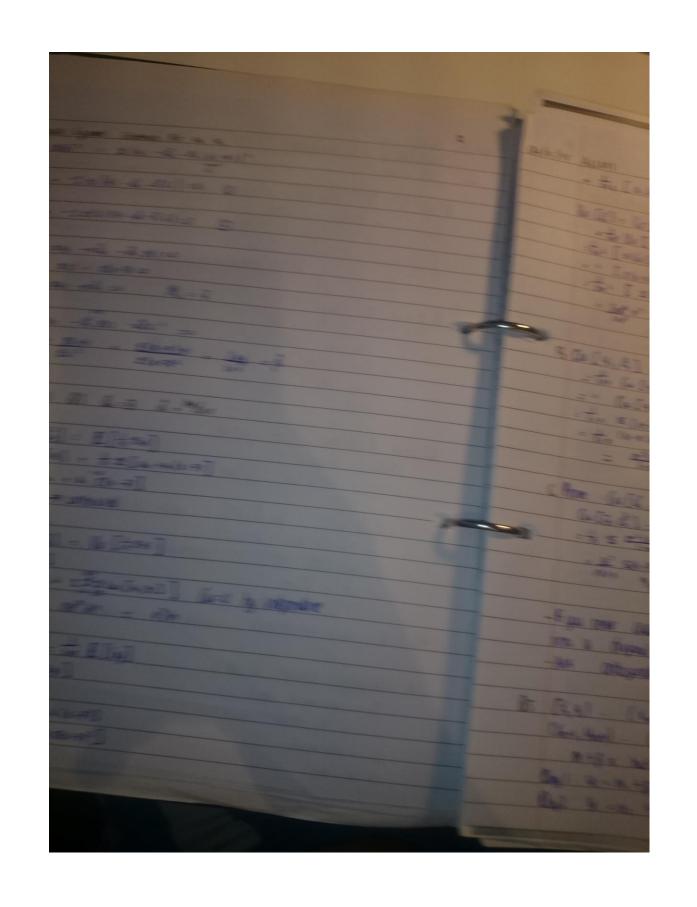


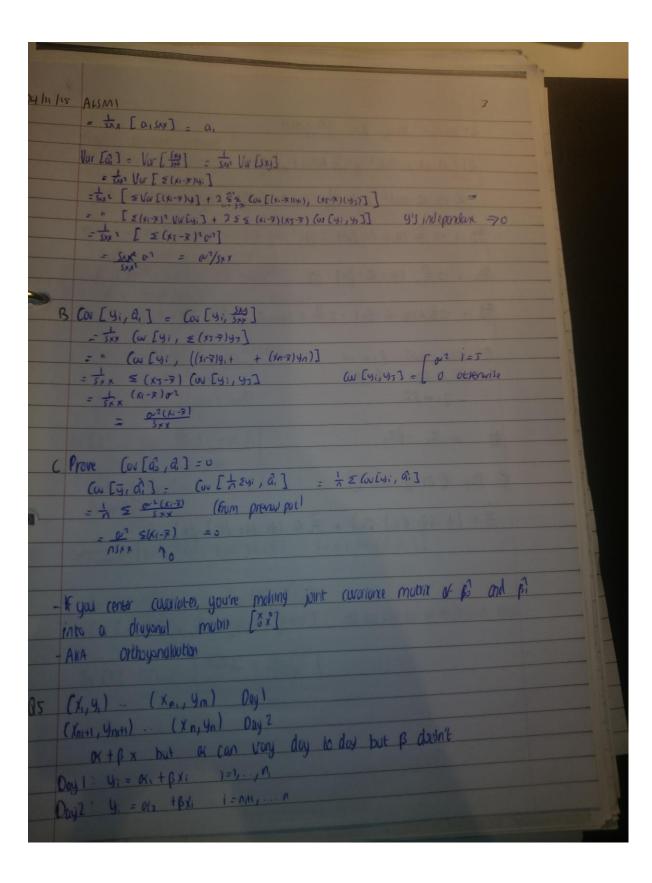


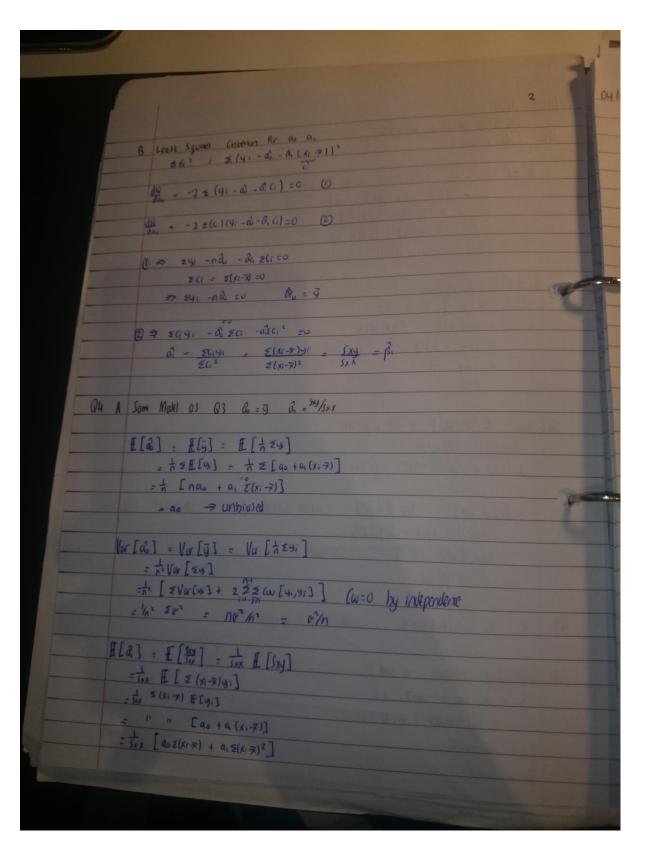
4/0 (ou - lak ord check it o beke as and of [ (w[do, a.]: (w[g, 54/2)] = NSIX (OU (EY), SXY] = 15xx (a) [ 4,+4,+.49,, 54] This [ (a) [4, 1/4] + (a) [4, 5/4]+ ... (a) [4, 1/4]] = NSVE \( \( \left( \omega \) \( \sigma \) \ The = (al 4; (x1-7/4+ +(x1-7)(9i)+ ++ (xn-x)(4) will only how cov or 9: with 4i become of independence = tisxx & Cov [4; (xi-x/4) - NSX1 (XI-Y) COU (Y:, Y:]  $= \frac{O^2}{\sqrt{2}} \frac{\mathcal{E}(x-x)}{2xx} = 0$ hepetitin ? no relutaritip between as and as, i.e no relationship between mean (y) and estands. Does not guarante independence of variables Varioto) do NOT change together.



B Least Square (norm for as a,  $\pm 61^2 = \Sigma (41 - 65 - 61)^2$ du = -2 = (41 - a3 - a, a) =0 (1) da - - 2 z(ci)(4: -2-0, ci)=0 (2) (1) => zy1 -nd -21 z(1=0 **Σ**(i = Σ(xi-π) =υ => Ey: -na =0 Ro = ÿ (2) = E(14) - a E(1 -a)E(1 =0  $\frac{G_1}{2G_1^2} = \frac{\mathcal{E}(x_1 - x_1)y_1}{\mathcal{E}(x_1^2 - x_1)^2} = \frac{\mathcal{E}(x_2 - x_1)y_1}{\mathcal{E}(x_1 - x_1)^2} = \frac{\mathcal{E}(x_1 - x_1)y_1}{\mathcal{E}(x_1 - x_1)^2} = \frac{\mathcal{$ Q4 A Same Model as Q3 2 = 9 2 = 549/5xx #[ab] = #[g] = #[h Eyi] = # = E[4,] = # = [a0 + a1 (x; 7)] = + [nas + a, E(x, ->)] = ao = unbiala Vor [a] = Vor [y] = Vor [th Eyi] = 1 Var [ = 4] = 1 [ = Var ( w ) + 2 = = con [ 41, 45] [ (w=0 by independence = 1/2 200 = NOThe = 01/11 FLa,] = F[sxy] = IXX F[sxy] = 5xx E[ E(x1-7/4)] = INX S(XI-X) E EYI]  $= \frac{11}{5\times x} \left[ a_0 \geq (x_1 \neq x) + a_1 \geq (x_1 \neq x)^2 \right]$ 







LS contorion show \$ = \frac{\widthered{0}}{\widthered{0}} + \frac{\widthered{0}}{\widthered{0}} Q(Q,Q,p) = 5812 = 2812 + 2812 = 2 (4: -a: -px) = + = (4: -a: -px) da = -2 € (y; -a, -βx) () 10 = -2 = (4: - 32 - 3x) (2) 10 - -2 xilyi-a, -3x) -2 = xilyi-a, -3xl 3 0 = 41 - 11,00 - BEX =0 0 = 241 - BEX - 9 = 48X 6) = 92 - BX2 10 Plug ( ord () Into (3) = xi (yi - (yi - \beta xi) - \beta xi) + \frac{2}{1200} xi (yi - (\beta z - \beta xi) - \beta xi) = 0

