1. E(e)=E(Y-E(Y|X))=E(Y)-E(E(Y|X))=E(Y)-E(Y)=0. This shows that the mean of the error in this model is 0. The covariance of X and e is  
   . We know that E(e)=0 We also know that since the mean of e is 0 the sum of all e’s is 0. This means the covariance is 0 and therefore the correlation between e and X is 0. This shows that they are uncorrelated.
2. We know that the Maximum Likelihood Estimator is actually equal to the NLSE. With this in mind the E(∂Q(B)/∂B) we expect it to be ∂Q(B)/∂B where we plug in the “most likely estimator” for B. This just so happens to be the NLSE and we know from the definition of a least squares estimator that ∂Q(BLS­)/∂B=0. Therefore, E(∂Q(B)/∂B)= ∂Q(BLS­)/∂B=0
3. Question 3
   1. The partial derivative with respect to B1 is. The partial derivative with respect to B2 is.
   2. The second order partial derivatives are and for B1 and B2 respectively.
   3. The score matrix . The heissan matrix is then.
   4. We will choose an initial value for B0 and an error d=10-5. For k=0,1,2… we can find Bk+1

. We will stop once ||Bk+1-Bk||<=d and we set our estimator Bhat =Bk+1

* 1. We will choose an initial value for B0 and set our error tolerance to d=10-5. For k=0,1,2… we calculate Bk+1=Bk.   
     When ||Bk+1-Bk||<=d we say our estimator Bhat=Bk+1.