Problem 4.2.

1. u is selected to be 2 by AIC, BIC and LRT at $\alpha = 0.05$. But since there are 2 predictors only, building an envelope with $u = 2$ will be same as fitting an OLS model. For this reason we decide to fix $u = 1$.
2. The OLS and envelope estimators of the coefficient vector are close to each other.
3. The envelope model gives gains in SE for all components of the coefficient matrix.
4. We have $\hat{\Gamma} = (-0.96, 0.29)$, which means Hemoglobin level and RBC count has the same type of effect on the hematocrit, and the effect due to Hemoglobin is larger than that due to RBC count.
5. The excluded point has actual hematocrit value 59.7, while its predicted hematocrit from the above envelope model is 48.9. This gives a *(y-yhat)/Std Error* = 1.71, will not be significant at 95% level for a t-test. Thus we conclude that the point is not influential.

Codes:

```
load AISxenv.txt
Y = AISxenv(:,1);
X = AISxenv(:,2:3);
[~,idx] = max(Y); Y(idx) = []; X(idx,:) = [];
maxdata = AISxenv(idx,:)

% select envelope dimension
u = modelselectaic(X, Y, 'xenv')
u = modelselectbic(X, Y, 'xenv')
alpha = 0.01;
u = modelselectlrt(X, Y, alpha,'xenv')

% Select u=1 since 2 is the full predictor dimension
u = 1; n = length(Y);

% Doing OLS and envelope fits, comparing models
envX = xenv(X, Y, u); olsX = fit_OLS(X,Y);
[olsX.betaOLS' envX.beta envX.asySE sqrt(olsX.n)*envX.beta./envX.asySE
envX.ratio envX.Gamma]

% Checking excluded point

MSE = sum((Y - X*envX.beta).^2)/n
maxpred = (maxdata(1)-maxdata(2:3)*envX.beta)^2
```