LAB 4 MATH 463

Suppose that

$$g(t) = \begin{cases} 0 & \text{if } t \leq 30 \\ \beta_1(t-30) & \text{if } 30 < t \leq 60, \\ \beta_1 30 + \beta_2(t-60) & \text{if } t > 60\,. \end{cases}$$

The model for the data is that

$$Y_i = g(a_i) + \epsilon_i\,. \tag{1}$$

The parameters β_i for $i = 1, 2$ are unknown.

The data for this problem is given at

`http://www.uoregon.edu/~dlevin/DATA/lab4data.csv`

Using indicator variables (i.e., variables which take on only the value 0 or 1), rewrite the model equation so that it is a multiple linear regression model.

Test the hypothesis that $\beta_1 = \beta_2$. Note that

$$\text{Var}(\hat{\beta}_1 - \hat{\beta}_2) = \text{Var}(\hat{\beta}_1) + \text{Var}(\hat{\beta}_2) - 2\text{Cov}(\hat{\beta}_1, \hat{\beta}_2)$$

Thus you will need the covariance of $\hat{\beta}_1$ with $\hat{\beta}_2$ to test the hypothesis that $\beta_1 = \beta_2$. You can obtain the unscaled covariances of estimated coefficients in a linear model `model` by `summary(model)$cov.unscaled`. Unscaled here means that the actual covariance will be σ^2 times the unscaled covariance.

You will need to create the indicator variables. To this end, functions such as

```
> x1 = function(x){
+      if((x>30)&&(x<=60)){v = 1}
+      else{v=0}
+      v
+ }
```

may be useful.

To apply `x1` to the vector `a`:

```
> xa = apply(matrix(a),1,x1)
```

Now generate (simulate) 100 new data points from the model (1) when the true values of β_1 and β_2 are 1 and 1.2, respectively. Assume that $\sigma = 30$. For the values of a, generate them uniformly in the interval $[0, 120]$.

The runif (random uniform) and rnorm (random normal) functions can be used:

```
> anew = runif(n=100,min=0,max=120)
> epsilon = rnorm(n=100,mean=0,sd=30)
```

Once you have the new simulated data, test the hypothesis that $\beta_1 = \beta_2$. Compare this with what you know about β_1 and β_2 in this case, since you do actually know the true values.