

REGRESSION IN-CLASS PROJECT - NCAA OVER/UNDER  
and TOTAL POINTS

1) For the population, subset the Total Points (Total Pts) and Over/Under (OvrUndr) to games where OT=0

2) Obtain the Parameters:  $\beta_0, \beta_1, \sigma^2$  for the model

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad Y \equiv \text{Total Pts} \quad X \equiv \text{Ovr Undr}$$

3) Obtain a plot of  $\varepsilon$  vs  $\beta_0 + \beta_1$  - Does constant variance seem reasonable?

4) Obtain a histogram of  $\varepsilon$  - Does normally distributed errors seem reasonable?

5) Set the seed to the combined first 3 digits of your UFID's

6) Obtain 100000 random samples of  $n=25$  games and for each sample, save:

$$\hat{\beta}_{n,1}, s^2, \text{ unique elements of } (X'X)^{-1}$$

7) Obtain the empirical means of:  $\hat{\beta}_{n,1}, s^2, (X'X)^{-1}$

Compare the means of  $\hat{\beta}_{n,1}$  and  $s^2$  w/  $\beta, \sigma^2$

8) Obtain the empirical variance-covariance matrix for  $\hat{\beta}_{n,1}$ , variance of  $s^2$ , compare with theoretical

9) Plot the regression slopes versus intercepts from the simulations