Math 584 Quiz 9 Spring 2021 Name_____

1) An overparameterized two way Anova model is $Y_{ijk} = \mu + \alpha_i + \beta_j + \tau_{ij} + \epsilon_{ijk}$ for $i = 1, ..., a$ and $j = 1, ..., b$ and $k = 1, ..., m$. Suppose $a = 2$, $b = 2$, and $m = 2$. Then

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix} = \boldsymbol{X} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \tau_{11} \\ \tau_{12} \\ \tau_{21} \\ \tau_{22} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \end{bmatrix}.$$

a) Give the matrix $\boldsymbol{X}$.

b) We can write the above model as $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. This model is **not full rank**. What is the projection matrix $\boldsymbol{P}$ (onto the column space of $\boldsymbol{X}$)? Hint: $\boldsymbol{X}^T\boldsymbol{X}$ is singular, so use the generalized inverse.

2) Suppose you are estimating the mean μ of losses with $T = \overline{X}$.
actual losses 14, 3, 5, 12, 20, 10, 9: $\overline{X} = 10.4286$,
a) Compute $T_1^*, ..., T_4^*$, where T_i^* is the sample mean of the i bootstrap sample.
bootstrap samples:

12, 3, 10, 14, 5, 9, 10:

10, 14, 5, 10, 10, 10, 9:

20, 5, 5, 3, 5, 20, 5:

12, 20, 5, 14, 12, 14, 20:

b) Now compute the bagging estimator which is the sample mean of the T_i^*: the bagging estimator $\overline{T}^* = \frac{1}{B}\sum_{i=1}^{B} T_i^*$ where $B = 4$ is the number of bootstrap samples.

3) The output below is for forward selection and I_{min} is the minimum C_p model. Here $Y =$ *height*, the constant $x_{i,1} \equiv 1$, $x_{i,2} =$ *height when sitting*, $x_{i,3} =$ *height when kneeling*, $x_{i,4} =$ *head length*, $x_{i,5} =$ *nasal breadth*, and $x_{i,6} =$ *span*.

```
           Estimate Std.Err 95% shorth CI
Intercept -42.4846 51.2863 [-192.281, 52.492]
X2          0              [   0.000,  0.268]
X3          1.1707  0.0598 [   0.992,  1.289]
X4          0              [   0.000,  0.840]
X5          0              [   0.000,  1.916]
X6          0.1467  0.0368 [  0.0747,  0.215]
  (Intercept)    a    b     c     d     e
1        TRUE FALSE TRUE FALSE FALSE FALSE
2        TRUE FALSE TRUE FALSE FALSE  TRUE
3        TRUE FALSE TRUE  TRUE FALSE  TRUE
4        TRUE FALSE TRUE  TRUE  TRUE  TRUE
5        TRUE  TRUE TRUE  TRUE  TRUE  TRUE
> tem2$cp
[1] 14.389492  0.792566  2.189839  4.024738  6.000000
```

What is the value of $C_p(I_{min})$ and what is $\hat{\boldsymbol{\beta}}_{I_{min},0}$?