Math 584 Quiz 9 Spring 2019 Name_____

1) The cell means model for the two way Anova model is $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$ for $i = 1, ..., a$ and $j = 1, ..., b$ and $k = 1, ..., m$. Suppose $a = 2$, $b = 2$, and $m = 2$. Then

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix} = \boldsymbol{X} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \end{bmatrix}.$$

want $X \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{pmatrix} = \begin{pmatrix} \mu_{11} \\ \mu_{11} \\ \mu_{12} \\ \mu_{12} \\ \mu_{21} \\ \mu_{21} \\ \mu_{22} \\ \mu_{22} \end{pmatrix} = E(Y)$

Give the matrix $\boldsymbol{X}$.

3.9 a)

$$\begin{bmatrix} 1000 \\ 1000 \\ 0100 \\ 0100 \\ 0010 \\ 0010 \\ 0001 \\ 0001 \end{bmatrix}$$

20

2) Suppose that a full rank cell means two way Anova model is written in matrix form as $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$.
What is the vector of residuals $\boldsymbol{e}$?

3.9 b) $(I - P)\underline{Y} = \underline{Y} - \hat{\underline{Y}}$

20

3) Suppose you are estimating the population median θ with the sample median $T = med(x)$. (Order the data and take the middle value.)

actual data: 14, 3, 5, 12, 20, 10, 9: $med(x) = 10$,

a) Compute T_1^*, T_2^*, T_3^*, where T_i^* is the sample median of the ith bootstrap sample.

bootstrap samples:

9, 10, 9, 12, 5, 14, 3 | 9

3, 9, 20, 10, 9, 5, 14 | 9

14, 12, 10, 20, 3, 3, 5 | 10

3 5 9 9 10 12 14

3 5 9 9 10 14 20

3 3 5 10 12 14 20

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

$\frac{9+9+10}{3} = \frac{28}{3} = 9.3333$

30

4) We will study the relationship between Y = the *number of women married to civilians* in the district with the predictors x_1 = constant, $x_2 = pop$ = the *population of the district in 1843*, $x_3 = mmen$ = the *number of married civilian men* in the district, $x_4 = mmilmen$ = *number of married men in the military* in the district, and $x_5 = milwmn$ = the *number of women married to husbands in the military* in the district. Sometimes the person conducting the survey would not count a spouse if the spouse was not at home.

```
            Estimate  Std.Err   95% shorth PI
Intercept   241.5445  190.7426  [ -218.40, 652.152]
pop=x2        0                 [ -0.0053,  0.0059]
mmen=x3       1.0010    0.0002  [  0.9687,  1.0310]
mmilmen=x4    0                 [ -0.5951,  7.9475]
milwmn=x5     0                 [ -8.7004,  0.5996]
out$cp
[1] -0.8268967  1.0151462  3.0029429  5.0000000
```

What is the value of $C_p(I_{min})$ and what is $\hat{\beta}_{I_{min},0}$? $C_p(I_{min}) = -0.8269$

$\hat{\beta}_{I_{min},0} = (241.5445, 0, 1.0010, 0, 0)'$

-7 if wrong

~15

~15

-8 if wrong

2

30