OldQuiz extra

Math 583 Exam 1 Fall 2017 Name____________________

1) The data below are a sorted residuals from a least squares regression where $n = 100$ and $p = 4$. Find shorth(97) of the residuals.

| number | 1 | 2 | 3 | 4 | ... | 97 | 98 | 99 | 100 |
|---|---|---|---|---|---|---|---|---|---|
| residual | -2.39 | -2.34 | -2.03 | -1.77 | ... | 1.76 | 1.81 | 1.83 | 2.16 |

1.76 + 2.39 = 4.15

1.81 + 2.34 = 4.15

1.83 + 2.03 = 3.86 ←

2.16 + 1.77 = 3.93

$\boxed{\text{Shorth}(97) = [-2.03, 1.83]}$

9

2) Suppose you are estimating the mean $\mu$ 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: 63/7 = 9

10, 14, 5, 10, 10, 10, 9: 68/7 = 9.7143

20, 5, 5, 3, 5, 20, 5: 63/7 = 9

12, 20, 5, 14, 12, 14, 20: 97/7 = 13.8571

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.

$$\frac{9 + 9.7143 + 9 + 13.8571}{4} = \frac{41.5714}{4}$$

$$= \boxed{10.39285}$$ 8 or 9



18

2) Suppose you are estimating the mean $\mu$ 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: 63/7 = 9

10, 14, 5, 10, 10, 10, 9: 68/7 = 9.7143

20, 5, 5, 3, 5, 20, 5: 63/7 = 9

12, 20, 5, 14, 12, 14, 20: 97/9 = 13.8571

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.

$$= \frac{9 + 9.7143 + 9 + 13.8571}{4} = \frac{41.5714}{4} = \boxed{10.3928}$$

-5

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*.

like a 4?3 Q11 problem but use AIC

```
                Estimate Std.Err 95% shorth CI
X1 ~~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 $\hat{\beta}_{I_{min},0}$? $Cp(I_{min}) = 0.792566$

$\hat{\beta}_{I_{min},0} = (-42.4846, 0, 1.1707, 0, 0, 0.1467)^T$

-5