Eli Pinkus
CS 156a
Set 7

```
E_out for the transform with k=3 is 0.42
E_val for the transform with k=3 is 0.3

E_out for the transform with k=4 is 0.416
E_val for the transform with k=4 is 0.5

E_out for the transform with k=5 is 0.188
E_val for the transform with k=5 is 0.2

E_out for the transform with k=6 is 0.084
E_val for the transform with k=6 is 0.0

E_out for the transform with k=7 is 0.072
E_val for the transform with k=7 is 0.1
```

The simulation returned the above output:
So we have the following answers for questions 1 and 2

1) D
2) E

After switching the validation set and the training set we get

```
E_out for the transform with k=3 is 0.396
E_val for the transform with k=3 is 0.28

E_out for the transform with k=4 is 0.388
E_val for the transform with k=4 is 0.36

E_out for the transform with k=5 is 0.284
E_val for the transform with k=5 is 0.2

E_out for the transform with k=6 is 0.192
E_val for the transform with k=6 is 0.08

E_out for the transform with k=7 is 0.196
E_val for the transform with k=7 is 0.12
```

So we have the following answers for questions 3,4, and 5

3) D
4) D
5) We have (.084,.196) so the answer is B

Eli Pinkus
CS 156a
Set 7

Problem 6

We must get a weighted average of all possible values of e_1 depending on how likely it is for $e_1 = \min(e_1, e_2)$ vs how likely it is for $e_2 = \min(e_1, e_2)$. We know that if $e_2 < e_1$, then since e_2 is uniformly distributed and less than e_1, we can say that it is uniformly distributed over $[0, e_1]$

This considered we arrive at the following integral which I will write in two separate forms that are equal:

$$\mathbb{E}[e] = \int_0^1 [\mathbb{P}[e_2 > e_1] \cdot e_1 + (1 - \mathbb{P}[e_2 > e_1]) \cdot \mathbb{E}(e_2)] de_1$$

Where we know for some e_1 we have:

$$\mathbb{P}[e_2 > e_1] = 1 - e_1$$

And

$$1 - \mathbb{P}[e_2 > e_1] = \mathbb{P}[e_1 > e_2] = e_1$$

and

$$\mathbb{E}[e_2] = \frac{e_1}{2}$$

so we have:

$$\mathbb{E}[e] = \int_{e_1=0}^{e_1=1} \left[(1 - e_1)e_1 + \frac{e_1^2}{2}\right] de_1 = \frac{1}{3} \approx 0.333 \ldots$$

$$\mathbb{E}(e) = \int_{e_1=0}^{e_1=1} \left[\int_{e_2=0}^{e_2=e_1} e_2 de_2 + \int_{e_2=e_1}^{e_2=1} e_1 de_2\right] de_1 = \frac{1}{3} \approx 0.333 \ldots$$

Which is closest to 0.4 so the answer is

Eli Pinkus
CS 156a
Set 7

Problem 7

Cross validation error:

$$E_{\mathrm{CV}} = \frac{1}{N}\sum_{n=1}^{N} e_n$$

and

$$E_{\mathrm{CV}} = \frac{1}{N}\sum_{n=1}^{N} e_n = \frac{1}{N}\sum_{n=1}^{N} (h_n(x) - y_n)^2$$

for the constant model we know that for any two points the hypothesis $h_1(x) = \frac{y_1+y_2}{2}$

so we have

$$E_{\mathrm{CV}} = \frac{1}{3}\left[\left(\frac{1}{2} - 0\right)^2 + (0-1)^2 + \left(\frac{1}{2} - 0\right)^2\right] = \frac{1}{4} + 1 + \frac{1}{4} = \frac{1}{3}\cdot\frac{3}{2} = \frac{1}{2}$$

For the linear model we have, given $(x_1, y_1), (x_2, y_2)$:

$$(h(x) - y_1) = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

$$\Rightarrow h(x) = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1$$

$$\Rightarrow h_n(x) = \frac{y_2 - y_1}{x_2 - x_1}\cdot x - \frac{y_2 - y_1}{x_2 - x_1}\cdot x_1 + y_1$$

and

$$E_{\mathrm{CV}} = \frac{1}{N}\sum_{n=1}^{N} e_n = \frac{1}{N}\sum_{n=1}^{N} (h_n(x) - y_n)^2$$

for leaving out $(-1,0)$ we have:

$$(x, y) = (-1,0)$$

$$(x_1, y_1) = (\rho, 1), \qquad (x_2, y_2) = (1,0)$$

Eli Pinkus
CS 156a
Set 7

$$e_1 = \left(\frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1 - y\right)^2 = \left(\frac{-1}{1-\rho}(-1-\rho) + 1\right)^2 = \left(\frac{1+\rho}{1-\rho} + 1\right)^2$$

for leaving out $(\rho, 1)$ we have:

$$(x, y) = (\rho, 1)$$
$$(x_1, y_1) = (-1,0), \quad (x_2, y_2) = (1,0)$$

$$e_2 = \left(\frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1 - y\right)^2 = \left((0+0) - 1\right)^2 = 1$$

for leaving out $(1,0)$ we have:

$$(x, y) = (1,0)$$
$$(x_1, y_1) = (-1,0), \quad (x_2, y_2) = (\rho, 1)$$
$$e_2$$

$$e_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1 - y\right)^2 = \left(\frac{1-0}{\rho+1}(1+1) + 0 - 0\right)^2 = \left(\frac{2}{\rho+1}\right)^2$$

so we want:

$$E_{\mathrm{CV}}(h_0) = E_{\mathrm{CV}}(h_1)$$

$$\frac{1}{3}\left[\left(\frac{2}{\rho+1}\right)^2 + 1 + \left(\frac{1+\rho}{1-\rho} + 1\right)^2\right] = \frac{1}{2}$$

$$\Rightarrow \rho = \sqrt{9 + 4\sqrt{6}}$$

by computer algebra system

So the answer is C

Eli Pinkus
CS 156a
Set 7

Question 8

The simulation returned SVM performing better 62.5% of the time so the answer is C

Question 9

The simulation returned SVM performing better 59.7% of the time so the answer is D

Question 10

The simulation returned an average number of support vectors 2.998 so the answer is B