

1. (b) See python figures
2. (d)  $c_{\min}$  is the only one where the samples are not independent.
3. (e)

$$P(y = f(x) \neq h(x)) = \lambda\mu$$

$$P(y \neq f(x) = h(x)) = (1 - \lambda)(1 - \mu)$$

So

$$P(y \neq h(x)) = \lambda\mu + (1 - \lambda)(1 - \mu)$$

4. (b) Collect all  $\mu$  terms and set the coefficient to zero.

$$\lambda\mu + 1 + \lambda\mu - \lambda - \mu$$

$$(2\lambda - 1)\mu = 0 \implies \lambda = \frac{1}{2}$$

5. (c)
6. (c)
7. (a)
8. (d) This also follows from symmetry since averaging out the linear regression weights implies that the line goes through the middle of the sample space implying that half the points are on the wrong side.
9. (a)

```
const    -1.050032
x         -0.067213
y         -0.096409
xy        -0.011164
x2         1.498617
y2         1.637846
```

10. (b)