1 point

1. If the features of Model 1 are a strict subset of those in Model 2, the TRAINING error of the two models can **never** be the same.

- ○ True
- ◉ False

---

1 point

2. If the features of Model 1 are a strict subset of those in Model 2, which model will USUALLY have lowest TRAINING error?

- ○ Model 1
- ◉ Model 2
- ○ It's impossible to tell with only this information

---

1 point

3. If the features of Model 1 are a strict subset of those in Model 2. which model will USUALLY have lowest TEST error?

- ○ Model 1
- ○ Model 2
- ◉ It's impossible to tell with only this information

---

1 point

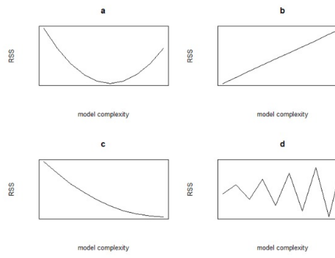
4. If the features of Model 1 are a strict subset of those in Model 2, which model will USUALLY have lower BIAS?

- ○ Model 1
- ◉ Model 2
- ○ It's impossible to tell with only this information

---

1 point

5. Which of the following plots of model complexity vs. RSS is most likely from TRAINING data (for a fixed data set)?

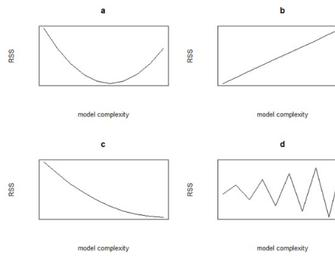


- ○ a
- ○ b
- ◉ c
- ○ d

---

1 point

6. Which of the following plots of model complexity vs. RSS is most likely from TEST data (for a fixed data set)?



- ◉ a
- ○ b
- ○ c
- ○ d

---

1 point

7. It is **always** optimal to add more features to a regression model.

- ○ True
- ◉ False

---

1 point

8. A simple model with few parameters is most likely to suffer from:

- ◉ High Bias
- ○ High Variance

---

1 point

9. A complex model with many parameters is most likely to suffer from:

- ○ High Bias
- ◉ High Variance

---

1 point

10. A model with many parameters that fits training data very well but does poorly on test data is considered to be

- ○ accurate
- ○ biased
- ◉ overfitted
- ○ poorly estimated

---

1 point

11. A common process for selecting a parameter like the optimal polynomial degree is:

- ○ Bootstrapping
- ○ Model estimation
- ○ Multiple regression
- ○ Minimizing test error
- ◉ Minimizing validation error

---

1 point

12. Selecting model complexity on test data (choose all that apply):

- ☐ Allows you to avoid issues of overfitting to training data
- ☑ Provides an overly optimistic assessment of performance of the resulting model
- ☐ Is computationally inefficient
- ☑ Should never be done

---

1 point

13. Which of the following statements is true (select all that apply): For a **fixed model complexity**, in the limit of an infinite amount of training data,

- ☐ The noise goes to 0
- ☐ Bias goes to 0
- ☑ Variance goes to 0
- ☐ Training error goes to 0
- ☐ Generalization error goes to 0