FALAK JAIN
22743504 52

1. D.
2. B
3. B
4. D
5. B
6. B
7. B
8. D
9. D
10. B
11. D
12. B
13. A
14. B
15. B
16. B

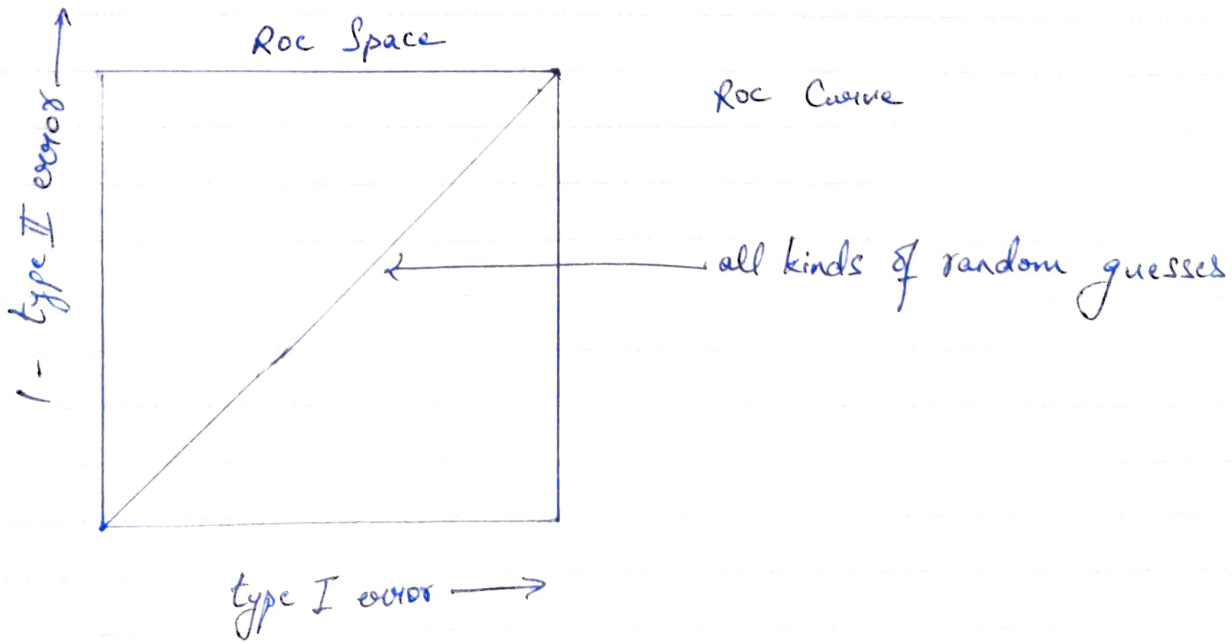
FALAK JAIN
22743504452

17. Steps of 9-nearest neighbors model

a. Iteratively select a point and compute distance between that point and every point in the dataset.
b. Find the 9 points in training set closest to the new point
c. Out of classes 1,2,3 find the most frequent class amongst those 9 points.
d. Assign the most frequent class to the new point

18. $$\frac{e^{\beta_0+3\beta_1+8\beta_2}}{1+e^{\beta_0+3\beta_1+8\beta_2}} \times \frac{e^{\beta_0+2\beta_1+4\beta_2}}{1+e^{\beta_0+2\beta_1+4\beta_2}} \times \frac{1}{1+e^{\beta_0+9\beta_1+8\beta_2}} \times \frac{1}{1+e^{\beta_0+10\beta_1+3\beta_2}}$$

19.



FALAK JAIN
22743504 52

20. $y = \beta_0 + \beta_1 x + \varepsilon \quad , \varepsilon \sim N(0, \sigma^2)$

A simple linear regression model has 3 parameters. ε

21.
- No, these errors are not satisfactory as the type I error is unacceptably large. In the type I error, class 0 is misclassified as 1.
- This means that many cancer cases are classified as normal which can have disastrous consequences which could be potentially fatal.
- I would increase the threshold to reduce the type I error

22.

$$d^2(0,1) = \frac{3}{1}(1-3)^2 = 12$$

$$d^2(0,2) = \frac{3}{1}(2-7)^2 = 75$$

$$d^2(0,3) = \frac{3}{2}((1-8)^2 + (2-9)^2) = 147$$

The closest 2 rows to row 0 are 1 and 2.

∴ Imputed value for row 0 blank is $\frac{3+5}{2} = 4$.