

## DM-Spring-2020-Q3-Grade

53.85% (7/13)

- ✓ 1. Logistic Regression Model is used to describe
- ☒ A Relationship between one categorical dependent variable and one or more (any) explanatory variables
  - ☐ B Relationship between one numeric dependent variable and one or more (any) explanatory variables
  - ☐ C Relationship between one categorical dependent variable and one explanatory variable
  - ☐ D Relationship between one categorical dependent variable and one or more numeric explanatory variables
  - ☐ E I do not know
- ✗ 2. Why Linear Regression cannot be used to predict the binary response variable?
- ☐ A Some of the estimates might be outside the [0,1] interval
  - ☐ B Coefficients of linear regression models do not exist
  - ☐ C There will be the multicollinearity
  - ☒ D All of the variants
  - ☐ E I do not know
- ✓ 3. The most common approach to estimate coefficients of logistic regression is
- ☒ A The Maximum Likelihood
  - ☐ B Ordinary Least Squares
  - ☐ C Generalized Method of Moments
  - ☐ D I do not know
- ✓ 4. The model of Logistic Regression is
- ☐ A  $\ln(\lambda) = e^{(xb)} / (1 + e^{(xb)})$
  - ☐ B  $\ln(y) = e^{(xb)} / (1 + e^{(xb)})$
  - ☒ C  $\Pr(y=1) = e^{(xb)} / (1 + e^{(xb)})$
  - ☐ D I do not know

✗ 5. We can estimate Logistic Regression in R using the function

- ☐ A `lm()`
- ☐ B `glm()`
- ☐ C `flm()`
- ☒ D `logit()`
- ☐ E I do not know

✗ 6. Which one of these is the correct interpretation of the coefficient of Logistic Regression?

- ☐ A For a 1-unit increase in X, we expect a b1 unit increase in Y.
- ☒ B For a 1-unit increase in X, we expect b1 percentage increase in Y.
- ☐ C For a 1-percentage increase in X, we expect b1 percentage increase in Y.
- ☐ D Increasing X by one unit changes the log odds by b1
- ☐ E I do not know

✗ 7. Logistic Regression cannot be used to model the response variable which

- ☐ A has two categories
- ☒ B has more than two categories
- ☐ C is ordinal
- ☐ D is numeric
- ☐ E I do not know

✓ 8. Accuracy =

- ☒ A  $(TP+TN)/Total$
- ☐ B  $TP/(TP+FN)$
- ☐ C  $TN/(TN+FP)$
- ☐ D I do not know

|        |              | Predicted    |              |
|--------|--------------|--------------|--------------|
|        |              | Negative (0) | Positive (1) |
| Actual | Negative (0) | TN           | FP           |
|        | Positive (1) | FN           | TP           |

✓ 9. Sensitivity =

- ☐ A  $(TP+TN)/Total$
- ☒ B  $TP/(TP+FN)$
- ☐ C  $TN/(TN+FP)$
- ☐ D I do not know

|        |              | Predicted    |              |
|--------|--------------|--------------|--------------|
|        |              | Negative (0) | Positive (1) |
| Actual | Negative (0) | TN           | FP           |
|        | Positive (1) | FN           | TP           |

✗ 10. Your lecturer decided that you are cheating while you are not. It is

- ☐ A Type 1 error (false positive)
- ☒ B Type 2 error (false negative)
- ☐ C I do not know

✓ 11. Suppose the data with the number of observations equals to 142, where 89 observations belong to class 1, and another part to 0. Let the level 1 is the positive case. We performed the logit model and obtained the accuracy = 60%. Does the model have a high predictive power?

- ☐ A Yes, because it is more than the non-information rate
- ☒ B No, because it is less than the non-information rate
- ☐ C Yes, because it is less than the non-information rate
- ☐ D No, as a result of other reasons.
- ☐ E The non-information rate? What is it?

✓ 12. Is Logistic regression a supervised machine learning algorithm?

- ☒ A Yes
- ☐ B No
- ☐ C I do not know

✗ 13. By using MLE for estimating the coefficient in the Logistic Regression model

- ☒ A we can obtain the unique formula for coefficients
- ☐ B we can obtain the unique formula for coefficients only for 1-D case
- ☐ C we cannot obtain the unique formula for coefficients
- ☐ D I do not know