Started on	Sunday, January 26, 2020, 2:50 PM
State	Finished
Completed on	Sunday, January 26, 2020, 2:54 PM
Time taken	4 mins 18 secs
Grade	**25.00** out of 25.00 (**100**%)

Question 1

Correct

5.00 points out of 5.00

Select all the equations that are valid forms of a linear regression model of predictors X1 and X2.

Select one or more:

- [ ] 1. None of the above
- [x] 2. $\beta_0 + \beta_1 * X_1 * X_2$ ✔
- [x] 3. $\beta_0 + \beta_1 * X_1^2$ ✔
- [ ] 4. $\beta_0 + \beta_1 * exp^{X_1/\beta_2}$
- [x] 5. $\beta_0 + \beta_1 * log(X_1)$ ✔
- [x] 6. $\beta_0 + \beta_1 * X_1 + \beta_2 * X_2^3$ ✔

Your answer is correct.

The correct answers are:

$\beta_0 + \beta_1 * X_1 * X_2$

,

$\beta_0 + \beta_1 * X_1^2$

,

$\beta_0 + \beta_1 * log(X_1)$

,

$\beta_0 + \beta_1 * X_1 + \beta_2 * X_2^3$

Question 2

Correct

10.00 points out of 10.00

Select all the assumptions that hold true for a linear regression model.

Select one or more:

- [x] 1. The response variable must be continuous. ✔
- [ ] 2. The predictor variables must be continuous.
- [x] 3. The predictor variables can be both categorical and continuous. ✔
- [x] 4. The response variable must be normally distributed. ✔
- [ ] 5. The predictor variables must follow the Gaussian distribution.
- [x] 6. Distribution of the response variable is a member of the exponential distribution family. ✔
- [ ] 7. In a linear regression model, the predictor variables can not be high-degree polynomial functions (e.g., squared, cubed).
- [x] 8. Observations used for deriving a linear regression model must be i.i.d., i.e., independent, identically distributed. ✔
- [x] 9. The residuals of the linear regression model must come from a normal distribution. ✔
- [ ] 10. The word "linear" in the linear regression means that the predictor variables can not be introduced as non-linear functions, such as polynomial, exponential, trigonometric, etc.
- [ ] 11. The errors/residuals of the linear regression model must be heteroscedastic.
- [ ] 12. A linear regression model is always sensitive to outliers.
- [x] 13. The errors/residuals of the linear regression model must be homoscedastic. ✔
- [x] 14. There are no distributional assumptions about the predictor variables. ✔
- [ ] 15. To ensure linear relationship with the response variable, the predictor variables must be multicollinear.
- [x] 16. A linear regression model of one predictor variable (X) may have the shape of a parabola. ✔

Your answer is correct.

The correct answers are: The response variable must be continuous., The predictor variables can be both categorical and continuous., The response variable must be normally distributed., Distribution of the response variable is a member of the exponential distribution family., Observations used for deriving a linear regression model must be i.i.d., i.e., independent, identically distributed., The residuals of the linear regression model must come from a normal distribution., The errors/residuals of the linear regression model must be homoscedastic., There are no distributional assumptions about the predictor variables. , A linear regression model of one predictor variable (X) may have the shape of a parabola.

Question 3

Correct

9.00 points out of 9.00

Select the proper regression model to build for the given problem.

1. What is the relationship between the number of calories burned and the duration of exercise (minutes), percentage of time spent at the target heart rate, average speed (mph), age (years), gender, and body mass index (BMI)?	Linear ✔
1. What is the expected number of credit cards a person may have, given his/her income?	Poisson ✔
What factors account for interstate differences in the price of beer?	Linear ✔
What is the number of people in line in front of you at the grocery store given the number of items currently offered at a special discounted price, information on special events (e.g., a holiday, a big sporting event) a few days away, etc.	Poisson ✔
What personal, demographic, and relationship variables predict marital infidelity (affair/no affair)?	Logistic ✔
Which qualities of an educational environment are most strongly related to higher educational achievement scores?	Linear ✔
How many people will visit your website every day between 9:00 AM and 10 AM?	Poisson ✔
What is the probability that the newborn baby will be at-risk given mother's pregnancy rate, the number of prenatal visits, mother's health conditions (diabetic, hypertension, seizures, etc), mother's habits (smoking, caffeine)?	Logistic ✔
What is the probability that the flight will be delayed given a flight's origin and destination, the time of year, the weather, and the air carrier?	Logistic ✔

Your answer is correct.

The correct answer is:

1. What is the relationship between the number of calories burned and the duration of exercise (minutes), percentage of time spent at the target heart rate, average speed (mph), age (years), gender, and body mass index (BMI)? → Linear,
1. What is the expected number of credit cards a person may have, given his/her income? → Poisson, What factors account for interstate differences in the price of beer? → Linear, What is the number of people in line in front of you at the grocery store given the number of items currently offered at a special discounted price, information on special events (e.g., a holiday, a big sporting event) a few days away, etc. → Poisson, What personal, demographic, and relationship variables predict marital infidelity (affair/no affair)? → Logistic, Which qualities of an educational environment are most strongly related to higher educational achievement scores? → Linear, How many people will visit your website every day between 9:00 AM and 10 AM? → Poisson, What is the probability that the newborn baby will be at-risk given mother's pregnancy rate, the number of prenatal visits, mother's health conditions (diabetic, hypertension, seizures, etc), mother's habits (smoking, caffeine)? → Logistic, What is the probability that the flight will be delayed given a flight's origin and destination, the time of year, the weather, and the air carrier? → Logistic

Question 4

Correct

1.00 points out of 1.00

Given the probability p, consider a special form of the sigmoid function, called the logit function defined as the natural logarithm of the odds: $logit(p) = ln(p/1-p)$. If the value of the logit function is 0, then what is the value of the probability p?

Answer: 0.5 ✔

The correct answer is: 0.5

◄ (DUE: 02/13/2020): SUBMIT: PROJECT: Network Properties with Apache Spark

Jump to...

(DUE: 01/30/2020): SUBMIT: QUIZ: Generalized Linear Model: Intermediate ►