Status	Finished
Started	Thursday, 7 November 2024, 12:31 PM
Completed	Thursday, 7 November 2024, 12:39 PM
Duration	7 mins 52 secs
Grade	**10.00** out of 10.00 (**100**%)

Question 1

Correct

Mark 1.00 out of 1.00

In the context of logistic regression, which of the following statements is true regarding the logit-link function?

- a. The logit of the probability p is given by $\log\left(\frac{p}{1-p}\right) = x^T\beta$.
- b. The logit function transforms p to a value between -1 and 1.
- c. The probability p of an event occurring is given by $p = x^T\beta$.
- d. Logistic regression does not use the concept of maximum likelihood estimation.

Your answer is correct.

The correct answer is: The logit of the probability p is given by $\log\left(\frac{p}{1-p}\right) = x^T\beta$.

Question 2

Correct

Mark 1.00 out of 1.00

Which of the following models uses the cumulative distribution function of the standard normal distribution in its link function?

- a. Decision Tree
- b. Logistic Regression with probit-link
- c. Gaussian Process Regression
- d. Linear Regression
- e. Random Forest
- f. Logistic Regression with logit-link

Your answer is correct.

The correct answer is: Logistic Regression with probit-link

Question 3

Correct

Mark 1.00 out of 1.00

In a logistic regression model, let $\{y_i, x_i\}_{i=1}^{n}$ represent n independent observations, where y_i is a binary outcome. The likelihood function for a single observation (y_i, x_i) is given by:

$f(y_i, x_i|\beta) = p_i^{y_i}(1 - p_i)^{1-y_i}$

where $p_i = \frac{\exp(x_i^T\beta)}{1+\exp(x_i^T\beta)}$. Which of the following expressions represents the joint likelihood $L(\beta; y, X)$ for the entire sample?

- a. $\prod_{i=1}^{n} \left[p_i^{y_i}(1 - p_i)^{1-y_i}\right]$
- b. $\prod_{i=1}^{n} [p_i + (1 - p_i)]$
- c. $\sum_{i=1}^{n} \left[\log(p_i^{y_i}) + \log((1 - p_i)^{1-y_i})\right]$
- d. $\sum_{i=1}^{n} \left[p_i^{y_i} + (1 - p_i)^{1-y_i}\right]$

Your answer is correct.

The correct answer is: $\prod_{i=1}^{n} \left[p_i^{y_i}(1 - p_i)^{1-y_i}\right]$

Question 4

Correct

Mark 1.00 out of 1.00

In logistic regression, let the probability p_i for the i-th observation be given by:

$p_i = \frac{\exp(x_i^T\beta)}{1+\exp(x_i^T\beta)}$

If the log-likelihood function $\ln L(\beta; y, X)$ is maximized, the Hessian matrix of the log-likelihood function, evaluated at the maximum likelihood estimate $\hat{\beta}$, is typically:

- a. Positive definite
- b. Positive semi-definite
- c. Indefinite
- d. Negative definite

Your answer is correct.

The correct answer is:
Negative definite

Question 5

Correct

Mark 1.00 out of 1.00

In the context of logistic regression with logit-link, the observed Fisher Information matrix $I(\beta)$ is given by:

$$I(\beta) = -\frac{\partial^2 \ln L(\beta; y, X)}{\partial\beta\partial\beta^T}$$

Which of the following best describes the purpose of $I(\beta)$?

- a. It gives a measure of model fit.
- b. It provides an estimate of the variance-covariance matrix of $\hat{\beta}$ for large sample sizes.
- c. It provides an estimate of the variance-covariance matrix of $\hat{\beta}$ for small sample sizes.
- d. It represents the likelihood function evaluated at β.

Your answer is correct.

The correct answer is:
It provides an estimate of the variance-covariance matrix of $\hat{\beta}$ for large sample sizes.

Question 6

Correct

Mark 1.00 out of 1.00

In logistic regression, suppose the asymptotic distribution of the maximum likelihood estimate $\hat{\beta}$ is given by:

$$\sqrt{n}(\hat{\beta} - \beta) \sim N(0, \Omega)$$

Which of the following statements is true about the matrix Ω?

- a. It is the variance-covariance matrix of $\hat{\beta}$.
- b. It is the identity matrix.
- c. It is always positive definite, regardless of the data.
- d. It is a diagonal matrix with entries equal to $p_i(1 - p_i)$.

Your answer is correct.

The correct answer is: It is the variance-covariance matrix of $\hat{\beta}$.

Question 7

Correct

Mark 1.00 out of 1.00

In Linear Discriminant Analysis (LDA), suppose the class-conditional density $f_k(x)$ for class k is modeled as a multivariate Gaussian distribution:

$$f_k(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu_k)^T\Sigma^{-1}(x-\mu_k)\right)$$

If we assume $\Sigma_k = \Sigma$ for all classes k, which of the following represents the decision boundary between classes k and l?

- a. $(x-\mu_k)^T(\mu_k-\mu_l) = (x-\mu_l)^T(\mu_k-\mu_l)$
- b. $(x-\mu_k)^T\Sigma^{-1}(x-\mu_k) = (x-\mu_l)^T\Sigma^{-1}(x-\mu_l)$
- c. $x^T\Sigma^{-1}(\mu_k-\mu_l) = \frac{1}{2}(\mu_k+\mu_l)^T\Sigma^{-1}(\mu_k-\mu_l)$
- d. $x^T\Sigma(\mu_k-\mu_l) = \frac{1}{2}(\mu_k-\mu_l)^T\Sigma(\mu_k-\mu_l)$

Your answer is correct.

The correct answer is:
$x^T\Sigma^{-1}(\mu_k-\mu_l) = \frac{1}{2}(\mu_k+\mu_l)^T\Sigma^{-1}(\mu_k-\mu_l)$

Question 8

Correct

Mark 1.00 out of 1.00

In LDA, the discriminant function $\delta_k(x)$ for class k is given by:

$$\delta_k(x) = x^T\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k + \log\pi_k$$

Which of the following represents the decision rule $G(x)$ for classifying a new observation x?

- a. $G(x) = \arg\max_k \delta_k(x)$
- b. $G(x) = \arg\min_k \left(x^T\Sigma^{-1}\mu_k\right)$
- c. $G(x) = \arg\max_k \left(-\frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k\right)$
- d. $G(x) = \arg\min_k \delta_k(x)$

Your answer is correct.

The correct answer is: $G(x) = \arg\max_k \delta_k(x)$

Question 9

Correct

Mark 1.00 out of 1.00

In Quadratic Discriminant Analysis (QDA), where $\Sigma_k \neq \Sigma$ for different classes k, the discriminant function $\delta_k(x)$ for class k is given by:

$\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x - \mu_k)^T\Sigma_k^{-1}(x - \mu_k) + \log \pi_k$

What type of decision boundary is formed between any two classes k and l?

- a. Polynomial of order 3 and above
- b. Linear
- c. Circular
- d. Quadratic
- e. Linear Hyper-Plane

Your answer is correct.

The correct answer is: Quadratic

Question 10

Correct

Mark 1.00 out of 1.00

In LDA, suppose we estimate the parameters π_k, μ_k, and Σ from the training data. If there are f_k observations for class k in a total of n observations, what is the maximum likelihood estimate of the prior probability π_k for class k?

- a. $\pi_k = \frac{f_k}{2n}$
- b. $\pi_k = \frac{n}{f_k}$
- c. $\pi_k = \frac{n-f_k}{n}$
- d. $\pi_k = \frac{f_k}{n}$

Your answer is correct.

The correct answer is:

$\pi_k = \frac{f_k}{n}$