



**UNIVERSITY OF THE PHILIPPINES MINDANAO**  
College of Science and Mathematics  
Department of Mathematics, Physics, and Computer Science



**AMAT 132 (Introductory Forecasting) — Exercise 8: Binary Logistic Regression**  
Saint Joy A. Mandalinao

**General Objectives:** This exercise aims to help students:

1. Conduct a binary logistic regression given a dataset with dependent and independent variables.
2. Interpret numerical results summary of the analysis.
3. Generate a binary logistic regression equation.
4. Compute for the  $P(X=1)$  for each row using the generated binary logistic regression equation.

**Instructions:**

1. Load the provided dataset in R.
2. Variables: a. Dependent variable: DV  
b. Independent variables: IV\_1, IV\_2, IV\_3, IV\_4, IV\_5
3. Run the code below to conduct a binary logistic regression.
4. Interpret numerical results.
5. Generate a binary logistic regression equation for the  $P(X=1)$
6. Compute for the  $P(X=1)$  for each data row using the generated binary logistic regression equation

**RESULTS AND DISCUSSIONS**

The screenshot displays the R Studio environment. The script editor on the left contains the following R code:

```
# Step 1: Read dataset (adjust the file path as needed)
data <- read_csv("C:/Users/saint/Documents/AMAT132/AMAT 132 Exercise 8 Dataset.csv")

# Step 2: Select relevant columns
independent_vars <- c("IV_1", "IV_2", "IV_3", "IV_4", "IV_5")
data <- data %>% select(DV, all_of(independent_vars))

# Step 3: Ensure DV is numeric and binary
data$DV <- as.numeric(data$DV)
data$DV <- ifelse(data$DV == 1, 1, 0)

# Step 4: Confirm only 0s and 1s in DV
stopifnot(all(data$DV %in% c(0, 1)))
print(table(data$DV))

# Step 5: Compute class imbalance weight
var_weight <- floor(nrow(Filter(data, DV == 0)) / nrow(Filter(data, DV == 1)))
print(paste("weight (rounded down):", var_weight))

# Step 6: Fit logistic regression with weights
logistic_model <- glm(DV ~ ., data = data,
  family = binomial(link = "logit"),
  weights = var_weight)
```

The Environment pane on the right shows the following objects:

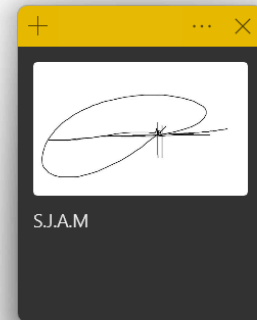
- data**: 271 obs. of 7 variables
- logistic\_model**: List of 30
- Values**:
  - conf\_matrix: 'table' int [1:2, 1:2] 173 42 16 40
  - independent\_vars: chr [1:5] "IV\_1" "IV\_2" "IV\_3" "IV\_4" "IV\_5"
  - predicted\_prob: Named num [1:271] 0.433 0.57 0.778 0.765 0.666 ...
  - var\_weight: 3

The Console pane at the bottom shows the output of the code execution, including the column specification of the dataset and the results of the model fitting process.

```
> # Step 1: Read dataset (adjust the file path as needed)
> data <- read_csv("C:/Users/saint/Documents/AMAT132/AMAT 132 Exercise 8 Dataset.csv")
Rows: 271 Columns: 7
# Column specification
Delimiter: ","
dbl (7): ID, DV, IV_1, IV_2, IV_3, IV_4, IV_5

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
> # Step 2: Select relevant columns
> independent_vars <- c("IV_1", "IV_2", "IV_3", "IV_4", "IV_5")
> data <- data %>% select(DV, all_of(independent_vars))
> # Step 3: Ensure DV is numeric and binary
> data$DV <- as.numeric(data$DV)
> data$DV <- ifelse(data$DV == 1, 1, 0)
> # Step 4: Confirm only 0s and 1s in DV
> stopifnot(all(data$DV %in% c(0, 1)))
> print(table(data$DV))

 0  1
215 56
> # Step 5: Compute class imbalance weight
> var_weight <- floor(nrow(filter(data, DV == 0)) / nrow(filter(data, DV == 1)))
> print(paste("Weight (rounded down):", var_weight))
```

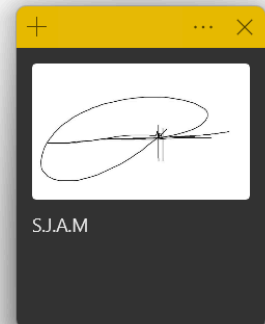


```
> # Step 5: Compute class imbalance weight
> var_weight <- floor(nrow(filter(data, DV == 0)) / nrow(filter(data, DV == 1)))
> print(paste("Weight (rounded down):", var_weight))
[1] "Weight (rounded down): 3"
> # Step 6: Fit logistic regression with weights
> logistic_model <- glm(DV ~ ., data = data,
+                       family = binomial(link = "logit"),
+                       weights = ifelse(data$DV == 1, var_weight, 1))
> # Step 7: Model summary
> summary(logistic_model)
```

Call:  
glm(formula = DV ~ ., family = binomial(link = "logit"), data = data,  
weights = ifelse(data\$DV == 1, var\_weight, 1))

Coefficients:

|             | Estimate   | Std. Error | z value | Pr(> z )     |
|-------------|------------|------------|---------|--------------|
| (Intercept) | 2.067e+00  | 4.148e-01  | 4.983   | 6.25e-07 *** |
| IV_1        | 1.040e-03  | 2.207e-04  | 4.714   | 2.43e-06 *** |
| IV_2        | -3.724e-02 | 1.560e-02  | -2.387  | 0.01701 *    |
| IV_3        | -3.135e-05 | 1.341e-03  | -0.023  | 0.98136      |
| IV_4        | -8.036e-05 | 1.299e-05  | -6.187  | 6.13e-10 *** |
| IV_5        | -2.801e-01 | 1.006e-01  | -2.783  | 0.00538 **   |



```
+ weights = ifelse(data$DV == 1, var_weight, 1))
> # Step 7: Model summary
> summary(logistic_model)
```

Call:  
glm(formula = DV ~ ., family = binomial(link = "logit"), data = data,  
weights = ifelse(data\$DV == 1, var\_weight, 1))

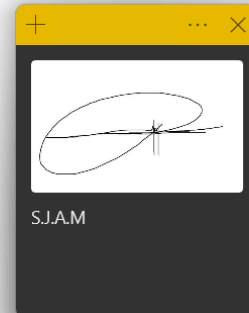
Coefficients:

|             | Estimate   | Std. Error | z value | Pr(> z )     |
|-------------|------------|------------|---------|--------------|
| (Intercept) | 2.067e+00  | 4.148e-01  | 4.983   | 6.25e-07 *** |
| IV_1        | 1.040e-03  | 2.207e-04  | 4.714   | 2.43e-06 *** |
| IV_2        | -3.724e-02 | 1.560e-02  | -2.387  | 0.01701 *    |
| IV_3        | -3.135e-05 | 1.341e-03  | -0.023  | 0.98136      |
| IV_4        | -8.036e-05 | 1.299e-05  | -6.187  | 6.13e-10 *** |
| IV_5        | -2.801e-01 | 1.006e-01  | -2.783  | 0.00538 **   |

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 525.17 on 270 degrees of freedom  
Residual deviance: 401.13 on 265 degrees of freedom  
AIC: 413.13



```

IV_4      -8.036e-05  1.299e-05  -6.187  6.13e-10 ***
IV_5      -2.801e-01  1.006e-01  -2.783  0.00538 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 525.17  on 270  degrees of freedom
Residual deviance: 401.13  on 265  degrees of freedom
AIC: 413.13

Number of Fisher Scoring iterations: 5

> # Step 8: Predict and classify
> predicted_prob <- predict(logistic_model, type = "response")
> data$Predicted <- ifelse(predicted_prob > 0.5, 1, 0)
> # Step 9: Confusion matrix
> conf_matrix <- table(Predicted = data$Predicted, Actual = data$DV)
> print(conf_matrix)
      Actual
Predicted 0   1
0      173  16
1       42  40

```



## Discussion Questions:

### 1. What variables have a significant relationship with the dependent variable? Interpret each coefficient.

The summary of a logistic regression model helps determine which independent variables are significantly related to the dependent variable. There are two ways to assess significance: First, by checking the p-values—values below 0.05 indicate statistical significance, while those above 0.05 suggest no significant relationship. Second, by looking at the stars next to the p-values—more stars imply stronger statistical significance. To interpret the relationship between each variable and the outcome, the coefficients are examined as follows:

#### IV\_1:

The coefficient for IV\_1 is 1.040e-03 (0.00104), and it is highly significant. This indicates that a one-unit increase in IV\_1 leads to a small increase of 0.00104 in the likelihood of the outcome being 1. Although the impact is small, it is still relevant due to its statistical significance.

#### IV\_2:

With a coefficient of -0.03724, IV\_2 is statistically significant. This means that a one-unit increase in IV\_2 results in a decrease of 0.03724 in the likelihood of the outcome being 1. While the effect size isn't very large, it remains meaningful because of its statistical significance.

#### IV\_3:

IV\_3 has a coefficient of -0.00003135 and is not statistically significant. This suggests that changes in IV\_3 do not meaningfully influence the outcome. As a result, IV\_3 likely does not contribute much to predicting whether the dependent variable equals 1.

#### IV\_4:

The coefficient for IV\_4 is -0.00008036, which is statistically significant. This means that a one-unit increase in IV\_4 slightly reduces the chance of the outcome being 1 by 0.00008036. Despite the small magnitude, the result is significant.

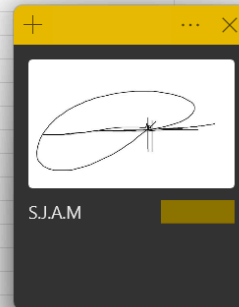
### IV\_5:

IV\_5 has a coefficient of -0.2801 and is statistically significant. This shows that a one-unit increase in IV\_5 leads to a more noticeable decrease of 0.2801 in the likelihood of the outcome being 1. Compared to the other variables, IV\_5 has a stronger influence on the outcome.

### Conclusion:

The logistic regression model indicates that IV\_1, IV\_2, IV\_4, and IV\_5 have statistically significant relationships with the dependent variable, each with varying degrees of impact. IV\_1 and IV\_4 show smaller but significant effects, while IV\_2 and IV\_5 display stronger impacts—with IV\_5 having the greatest effect. On the other hand, IV\_3 is not significant and likely has minimal influence on the outcome. Overall, IV\_2 and IV\_5 are the most influential predictors, whereas IV\_1 and IV\_4 offer minor but meaningful contributions.

| H2 : $\frac{1}{1 + \exp(-(\text{\$K\$1} + (\text{\$K\$2} * \text{C2}) + (\text{\$K\$3} * \text{D2}) + (\text{\$K\$4} * \text{E2}) + (\text{\$K\$5} * \text{F2}) + (\text{\$K\$6} * \text{G2})))}$ |    |    |      |      |      |       |      |                      |   |               |          |
|---|----|----|------|------|------|-------|------|----------------------|---|---------------|----------|
|   | A  | B  | C    | D    | E    | F     | G    | H                    | I | J             | K        |
| 1   | ID | DV | IV_1 | IV_2 | IV_3 | IV_4  | IV_5 | Computed Probability |   | Intercept ?0? | 2.067    |
| 2   | 1  | 0  | 175  | 12   | 68   | 12500 | 3.8  | 0.433178342          |   | ?1? for IV_1  | 0.00104  |
| 3   | 2  | 1  | 450  | 12   | 63   | 5400  | 4.9  | 0.569474586          |   | ?2? for IV_2  | -0.03724 |
| 4   | 3  | 1  | 450  | 12   | 2    | 6900  | 1    | 0.77790051           |   | ?3? for IV_3  | -3.1E-05 |
| 5   | 4  | 1  | 450  | 12   | 5    | 6400  | 1.4  | 0.765220814          |   | ?4? for IV_4  | -8E-05   |
| 6   | 5  | 1  | 450  | 12   | 71   | 6900  | 3    | 0.66621305           |   | ?5? for IV_5  | -0.2801  |
| 7   | 6  | 0  | 450  | 12   | 259  | 6400  | 5.9  | 0.478287014          |   |               |          |
| 8   | 7  | 0  | 450  | 12   | 122  | 15500 | 4.8  | 0.376185471          |   |               |          |
| 9   | 8  | 1  | 500  | 12   | 99   | 8000  | 4    | 0.592360308          |   |               |          |
| 10  | 9  | 1  | 500  | 12   | 71   | 8000  | 2.8  | 0.670562576          |   |               |          |
| 11  | 10 | 1  | 500  | 12   | 63   | 17000 | 3    | 0.482937078          |   |               |          |
| 12  | 11 | 1  | 500  | 12   | 132  | 22800 | 3.9  | 0.312464457          |   |               |          |
| 13  | 12 | 1  | 500  | 12   | 69   | 8000  | 2    | 0.718062946          |   |               |          |
| 14  | 13 | 1  | 500  | 12   | 136  | 27200 | 1.8  | 0.364908511          |   |               |          |
| 15  | 14 | 1  | 500  | 12   | 1    | 6500  | 2.1  | 0.736826434          |   |               |          |
| 16  | 15 | 1  | 500  | 12   | 10   | 8000  | 1.2  | 0.761479138          |   |               |          |
| 17  | 16 | 1  | 500  | 12   | 136  | 8000  | 0.5  | 0.794606537          |   |               |          |
| 18  | 17 | 1  | 500  | 12   | 160  | 8000  | 3    | 0.657442099          |   |               |          |
| 19  | 18 | 0  | 500  | 12   | 151  | 11600 | 3.9  | 0.527674837          |   |               |          |
| 20  | 19 | 0  | 550  | 12   | 168  | 9100  | 1.9  | 0.715730059          |   |               |          |
| 21  | 20 | 0  | 550  | 36   | 0    | 22300 | 3    | 0.208503769          |   |               |          |
| 22  | 21 | 0  | 550  | 12   | 39   | 27200 | 3    | 0.302547399          |   |               |          |
| 23  | 22 | 1  | 550  | 12   | 51   | 9100  | 1    | 0.764793606          |   |               |          |
| 24  | 23 | 1  | 550  | 12   | 59   | 9100  | 1    | 0.764748416          |   |               |          |
| 25  | 24 | 0  | 550  | 36   | 226  | 22650 | 3    | 0.202750681          |   |               |          |
| 26  | 25 | 0  | 550  | 12   | 121  | 13100 | 6    | 0.367026292          |   |               |          |
| 27  | 26 | 0  | 550  | 12   | 187  | 26350 | 5    | 0.208872058          |   |               |          |



|    | A  | B | C     | D  | E   | F     | G   | H           | I | J | K |
|----|----|---|-------|----|-----|-------|-----|-------------|---|---|---|
| 25 | 24 | 0 | 550   | 36 | 226 | 22650 | 3   | 0.202750681 |   |   |   |
| 26 | 25 | 0 | 550   | 12 | 121 | 13100 | 6   | 0.367026292 |   |   |   |
| 27 | 26 | 0 | 550   | 12 | 187 | 26350 | 5   | 0.208872058 |   |   |   |
| 28 | 27 | 0 | 550   | 36 | 92  | 28900 | 2   | 0.169787674 |   |   |   |
| 29 | 28 | 0 | 550   | 12 | 172 | 25950 | 5.8 | 0.178990375 |   |   |   |
| 30 | 29 | 0 | 550   | 12 | 252 | 45300 | 5.7 | 0.045102307 |   |   |   |
| 31 | 30 | 0 | 550   | 12 | 151 | 13100 | 5   | 0.433929291 |   |   |   |
| 32 | 31 | 0 | 550   | 12 | 186 | 13950 | 5.1 | 0.41017248  |   |   |   |
| 33 | 32 | 0 | 550   | 12 | 199 | 21200 | 5   | 0.285312442 |   |   |   |
| 34 | 33 | 0 | 550   | 12 | 158 | 14600 | 4.8 | 0.418100188 |   |   |   |
| 35 | 34 | 0 | 550   | 12 | 21  | 66900 | 3   | 0.017551243 |   |   |   |
| 36 | 35 | 0 | 550   | 12 | 161 | 9100  | 2.6 | 0.674262707 |   |   |   |
| 37 | 36 | 0 | 550   | 36 | 114 | 22300 | 2.9 | 0.212564231 |   |   |   |
| 38 | 37 | 0 | 550   | 12 | 56  | 13300 | 2   | 0.636770305 |   |   |   |
| 39 | 38 | 0 | 550   | 12 | 93  | 20000 | 2   | 0.505449734 |   |   |   |
| 40 | 39 | 0 | 550   | 36 | 119 | 28900 | 1.9 | 0.173650866 |   |   |   |
| 41 | 40 | 0 | 550   | 36 | 45  | 23000 | 3   | 0.199147186 |   |   |   |
| 42 | 41 | 0 | 550   | 36 | 121 | 9100  | 2.8 | 0.444974915 |   |   |   |
| 43 | 42 | 0 | 550   | 36 | 84  | 22300 | 3   | 0.208068821 |   |   |   |
| 44 | 43 | 0 | 550   | 36 | 1   | 22300 | 3   | 0.208498587 |   |   |   |
| 45 | 44 | 0 | 550   | 36 | 0   | 8600  | 1   | 0.581069175 |   |   |   |
| 46 | 45 | 0 | 600   | 12 | 290 | 22150 | 3   | 0.404858878 |   |   |   |
| 47 | 46 | 0 | 600   | 12 | 129 | 27100 | 5   | 0.20780843  |   |   |   |
| 48 | 47 | 0 | 634   | 12 | 131 | 13000 | 6.9 | 0.331368198 |   |   |   |
| 49 | 48 | 0 | 634   | 12 | 73  | 6000  | 3.8 | 0.674947412 |   |   |   |
| 50 | 49 | 0 | 637.5 | 12 | 7   | 8900  | 3   | 0.674233627 |   |   |   |
| 51 | 50 | 1 | 650   | 12 | 3   | 8800  | 1.8 | 0.747378192 |   |   |   |
| 52 | 51 | 1 | 650   | 12 | 143 | 22300 | 2   | 0.20626255  |   |   |   |

