

# Class Activity 25

Your name here

February 27 2024

## Group Activity 1

In this activity, we will calculate the probability of diabetes for a glucose level of 150 mg/dL using the logistic regression coefficients  $\beta_0 = -5.61$  and  $\beta_1 = 0.0392$ .

### a. Calculate Log Odds

First, calculate the log odds for a glucose level of 150 mg/dL.

### b. Convert Log Odds to Odds

### c. Convert Odds to Probability

Finally, convert the odds to probability.

The probability of having diabetes at a glucose level of 150 mg/dL is calculated to be

---

## Group Activity 2

- a. Let's fit the logistic regression model to predict the `diabetes` status given the `glucose` level.

```
set.seed(12345)
db_single <- db %>% select(_____, _____)
db_split <- initial_split(db_single, prop = 0.80)

# Create training data
db_train <- db_split %>% training()

# Create testing data
db_test <- db_split %>% testing()

fitted_logistic_model <- logistic_reg() %>% # Call the model function
  # Set the engine/family of the model
  set_engine() %>%
  # Set the mode
  set_mode() %>%
  # Fit the model
```

```

fit()

tidy(fitted_logistic_model)
Error: <text>:2:29: unexpected input
1: set.seed(12345)
2: db_single <- db %>% select(_

```

- b. We are interested in predicting the `diabetes` status of patients depending on the amount of `glucose`. Verify that the `glucose` value of 143.11 gives the probability of having `diabetes` as 1/2.

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

Answer:

- c. What value of glucose is needed to have a probability of diabetes of 0.75?

Answer:

- d. Make a classifier that classifies the diabetes status of new patients with a threshold of 0.5, i.e, a new patient is classified as negative if the estimated class probability is less than 0.5. Also, create a confusion matrix of the resulting predictions. Evaluate the model based on accuracy, sensitivity, specificity, and ppv.

```

# Prediction Probabilities

pred_prob <- predict(, new_data = , type = )
Error in UseMethod("predict"): no applicable method for 'predict' applied to an object of class "name"

db_results <- db_test %>% bind_cols(pred_prob) %>%
  mutate(.pred_class = make_two_class_pred(.pred_neg, levels(diabetes), threshold = .55)) %>%
  select(diabetes, glucose, contains(".pred"))
Error in eval(expr, envir, enclos): object 'db_test' not found

db_results %>%
  conf_mat(diabetes, .pred_class) %>%
  autoplot(type = "heatmap")
Error in eval(expr, envir, enclos): object 'db_results' not found

# Evaluating the model
eval_metrics <- metric_set(accuracy, sensitivity, specificity, ppv)

eval_metrics(data = db_results,
  truth = diabetes,
  estimate = .pred_class) %>% select(-2)
Error in `metric_set()`:
! Failed to compute `accuracy()`.
Caused by error:
! object 'db_results' not found

```

- e. Evaluate the performance of a diabetes prediction model at different classification thresholds and visualize how various metrics such as accuracy, sensitivity, and PPV change across these thresholds. Use a sequence of threshold values, apply each one to classify test data, calculate the performance metrics for each classification, and then create a line plot to illustrate the results.

```

# Step 1: Generate a sequence of thresholds
thresholds <-

# Step 2: Calculate metrics for each threshold using map
metrics_list <- map_df(, ~{
  db_results <- db_test %>%
    bind_cols(pred_prob) %>%
    mutate(.pred_class = make_two_class_pred(.pred_neg, levels(diabetes), threshold = )) %>%
    select(diabetes, glucose, contains(".pred"))

  metrics <- eval_metrics(data = db_results, truth = diabetes, estimate = .pred_class) %>%
    mutate(threshold = ) %>%
    select(-2)
  return(metrics)
})
Error in map_df(, ~{: argument ".x" is missing, with no default

# Step 3: Plot the metrics across thresholds
ggplot(metrics_list, aes(x = , y = , color = )) +
  geom_line() +
  labs(title = "Model Performance Metrics Across Thresholds",
       x = "Threshold",
       y = "Metric Value") +
  theme_minimal() +
  scale_color_viridis_d(begin = 0.2, end = 0.8, direction = 1, option = "C")
Error in eval(expr, envir, enclos): object 'metrics_list' not found

```