Week 6 in-class activities/Lab

Task 1a: Interpreting Logistic Regression model

Answer the following questions:
(you may use the provided "logistic regression" notebook and AI assistant.)

    a.   Thomas studied for two hours and did not attend the review session. What is his (1) log odds, (2) odds, and (3) likelihood of passing the exam?

```
# Thomas's situation
hours_studied_thomas = 2
review_session_thomas = 0

log_odds_thomas = log_odds(hours_studied_thomas, review_session_thomas)




odds_thomas = np.exp(log_odds_thomas)
probability_thomas = logistic_function(log_odds_thomas)

print("1. Log Odds:", log_odds_thomas)
print("2. Odds:", odds_thomas)
print("3. Likelihood of Passing:", probability_thomas)
```

#output
1.Log Odds: -1.4
2. Odds: 0.2465969639416065
3. Likelihood of Passing: 0.19781611144141825

b.If Thomas goes to the review session, what is the updated 1) log_odds, (2) odds, and (3) likelihood of passing the exam?

```
# Thomas's original situation
hours_studied_thomas = 2
review_session_thomas_original = 0

# Calculate the original log odds
log_odds_thomas_original = log_odds(hours_studied_thomas, review_session_thomas_original)

# Calculate the original odds
odds_thomas_original = np.exp(log_odds_thomas_original)
```

```
# Calculate the original likelihood of passing
probability_thomas_original = logistic_function(log_odds_thomas_original)

# Thomas's updated situation with a review session
review_session_thomas_updated = 1

# Calculate the updated log odds
log_odds_thomas_updated = log_odds(hours_studied_thomas, review_session_thomas_updated)

# Calculate the updated odds
odds_thomas_updated = np.exp(log_odds_thomas_updated)

# Calculate the updated likelihood of passing
probability_thomas_updated = logistic_function(log_odds_thomas_updated)

# Print the results
print("Original Log Odds:", log_odds_thomas_original)
print("Original Odds:", odds_thomas_original)
print("Original Likelihood of Passing:", probability_thomas_original)

print("Updated Log Odds:", log_odds_thomas_updated)
print("Updated Odds:", odds_thomas_updated)
print("Updated Likelihood of Passing:", probability_thomas_updated)
```

#output
Original Log Odds: -1.4
Original Odds: 0.2465969639416065
Original Likelihood of Passing: 0.19781611144141825
Updated Log Odds: 0.10000000000000009
Updated Odds: 1.1051709180756477
Updated Likelihood of Passing: 0.52497918747894

c.If Thomas studied more or less hours, would the answer change?
```
# Define a different number of hours studied (e.g., 3 hours)
hours_studied_thomas_updated = 3

# Calculate the updated log odds with the new hours studied
log_odds_thomas_updated = log_odds(hours_studied_thomas_updated,
review_session_thomas_updated)
```

```
# Calculate the updated odds
odds_thomas_updated = np.exp(log_odds_thomas_updated)

# Calculate the updated likelihood of passing
probability_thomas_updated = logistic_function(log_odds_thomas_updated)

# Print the results for the updated hours studied
print("Updated Hours Studied:", hours_studied_thomas_updated)
print("Updated Log Odds:", log_odds_thomas_updated)
print("Updated Odds:", odds_thomas_updated)
print("Updated Likelihood of Passing:", probability_thomas_updated)
```

#output
Updated Hours Studied: 3
Updated Log Odds: 0.9000000000000004
Updated Odds: 2.4596031111569507
Updated Likelihood of Passing: 0.710949502625004

In summary, changes in the number of hours studied directly impact the log odds, odds, and likelihood of passing according to the coefficient associated with Hours_Studied in the logistic regression model.

d.How would you interpret the coefficient of review_session (1.5) from the above experiment?
The coefficient of 'review_session' (1.5) in the logistic regression model indicates that attending a review session significantly increases the likelihood of passing the exam. Specifically, students who attend a review session are approximately 4.48 times more likely to pass compared to those who do not attend, when considering other factors in the model. This positive coefficient reflects the strong positive impact of attending a review session on exam success.

e.Using similar reasoning, how would you interpret the coefficient of hours_studied (0.8)
The coefficient of 'hours_studied' (0.8) in the logistic regression model indicates that studying more hours significantly increases the likelihood of passing the exam. Specifically, for every additional hour studied, the odds of passing the exam are about 2.22 times higher, assuming all other factors remain constant. This positive coefficient reflects the strong positive impact of studying time on exam success.

f.How would you interpret the intercept?
The intercept in your logistic regression model, which is -3, represents the baseline log odds of passing the exam when no review session is attended and no hours are studied. This means that, without any review session or studying, the odds of passing the exam are less than 1. The

intercept serves as a reference point for understanding how other factors, such as attending a review session and studying hours, influence the log odds and, consequently, the likelihood of passing the exam.

g.For someone who studied 8 hours, would you recommend him/her to attend the review session?

```
# Situation 1: 8 hours studied, no review session
hours_studied = 8
review_session = 0
log_odds_no_review = log_odds(hours_studied, review_session)
```

```
# Situation 2: 8 hours studied, attending a review session
review_session = 1
log_odds_review = log_odds(hours_studied, review_session)
```

```
# Calculate the probabilities
probability_no_review = logistic_function(log_odds_no_review)
probability_review = logistic_function(log_odds_review)
```

```
print("Probability of Passing without Review Session:", probability_no_review)
print("Probability of Passing with Review Session:", probability_review)
```
#output
Probability of Passing without Review Session: 0.9677045353015495
Probability of Passing with Review Session: 0.9926084586557181

In summary, students who have not dedicated as much time to studying or did not initially attend the review session are likely to benefit the most from participating in the review session. This session can provide them with an extra boost in their likelihood of passing the exam.

h.What type of students seems to benefit most from the review session?

In your logistic regression model, both 'hours_studied' and 'review_session' have positive coefficients, indicating that both studying more hours and attending the review session increase the log odds of passing the exam.

Students who have already studied a significant number of hours, particularly 8 or more, may benefit the most from the review session. This is because their higher study time, combined with the positive effect of the review session, leads to a substantial increase in the log odds of passing.

However, students who have studied fewer hours can still benefit from the review session, although the impact may not be as pronounced. The positive effect of the review session helps improve their chances of passing, but the lower study time may limit the extent of this improvement.

In summary, both high and low studiers can benefit from the review session. High studiers may see a more substantial increase in their chances of passing, while low studiers can still experience an improvement, though not as pronounced.