| | |
|---|---|
| **Started on** | Thursday, February 6, 2020, 5:43 PM |
| **State** | Finished |
| **Completed on** | Thursday, February 6, 2020, 5:48 PM |
| **Time taken** | 4 mins 37 secs |
| **Grade** | **30.00** out of 30.00 (**100**%) |

Question **1**

Correct

5.00 points out of 5.00

Suppose that a university only includes two Schools: The School of Computer Science and The Schools of Management. The total number of students is 1,100, with 100 students in Computer Science (CS). Based on some prior surveys, you know that about 80% students in the CS department wear jeans, yet only 16% Business students wear jeans.

You met Sam on the campus of this university, wearing jeans. Select the proper matches.

HINT: See slides #9 and 10 in Bayesian Inference Lecture

| | |
|---|---|
| The knowledge prior odds ratio is | 1:10 ✔ |
| The likelihood ratio is | 80:16 ✔ |
| The posterior odds ratio is | 1:2 ✔ |
| The probability that Sam wearing the jeans is from the Business school is | 0.667 ✔ |

Knowledge prior odds ratio:

P(CS) = 100/1100, P(Bus.)=1000/1100

P(CS)/P(Bus.) = 100/1000 = 1:10

Likelihood ratio:

P(Jeans | CS)=0.8 and P(Jeans | Bus.)=0.16

P(Jeans | CS)/P(Jeans | Bus.)=0.8:0.16 = 80:16

Posterior odds ratio:

P(CS | Jeans) ~ likelihood x priors = 0.8 x (100/1100) = 0.8/11

P(Bus. | Jeans) ~ likelihood x priors = 0.16 x (1000/1100) = 1.6/11

P(CS | Jeans):P(Bus. | Jeans) = 0.8:1.6 = 1:2

The probability that Sam wearing the jeans is from the Business school is

P(Bus. | Jeans) = (P(Jeans | Bus.) x P(Bus.))/P(Jeans)

= (0.16 x 1000/1100)/(240/1100)

= 0.667

The correct answer is: The knowledge prior odds ratio is → 1:10, The likelihood ratio is → 80:16, The posterior odds ratio is → 1:2, The probability that Sam wearing the jeans is from the Business school is → 0.667

Question 2

Correct

3.00 points out of 3.00

Select all that hold true when you apply the chain rule to the joint probability distribution P(A,B,C) of the three random variables:

Select one or more:

- a. P(A, B, C) = P(A) * P(B) * P(C)
- b. P(A, B, C) = P(A|B,C) * P(B|C) * P(C) ✔
- c. P(A, B, C) = P(A|B,C) * P(B|A,C) * P(C|A,B)
- d. P(A, B, C) = P(A) * P(B|A) * P(C|A,B) ✔
- e. P(A, B, C) = P(B|A,C) * P(C|A) * P(C)

Your answer is correct.

The correct answers are: P(A, B, C) = P(A|B,C) * P(B|C) * P(C), P(A, B, C) = P(A) * P(B|A) * P(C|A,B)

Question 3

Correct

4.00 points out of 4.00

Suppose that each of the two random variables A and B can only take on two values, 0 or 1.

Given the joint probability distribution P(A,B), select the correct formula for P(B):

Select one:

- a. P(B) = P(A=0, B) + P(A=1, B) ✔
- b. P(B) = P(A) * P(B | A)
- c. P(B) = P(A, B) * P(A)

Your answer is correct.

The correct answer is: P(B) = P(A=0, B) + P(A=1, B)

Question 4

Correct

4.00 points out of 4.00

Given the joint probability distribution P(A, B, C) of the three discrete random variables, select the correct formula for P(A):

Select one:

- a. $P(A) = \sum_B \sum_C P(A, B, C)$ ✔
- b. $P(A) = \sum_B \sum_C P(A|B, C)$
- c. $P(A) = \sum_B P(A|B) + \sum_C P(A|C)$

Your answer is correct.

The correct answer is: $P(A) = \sum_B \sum_C P(A, B, C)$

Question 5

Correct

4.00 points out of 4.00

Given the data in the Table below, compute **P ( Evade = NO | Status = Divorced )**. Round your answer to the two decimal points.

| *Tid* | Refund | Marital Status | Taxable Income | Evade |
|---|---|---|---|---|
| 1 | Yes | Single | 125K | No |
| 2 | No | Married | 100K | No |
| 3 | No | Single | 70K | No |
| 4 | Yes | Married | 120K | No |
| 5 | No | Divorced | 95K | Yes |
| 6 | No | Married | 60K | No |
| 7 | Yes | Divorced | 220K | No |
| 8 | No | Single | 85K | Yes |
| 9 | No | Married | 75K | No |
| 10 | No | Single | 90K | Yes |

Answer: 0.50 ✔

The correct answer is: 0.5

Question 6

Correct

5.00 points out of 5.00

Given the data in the Table below, compute **P ( Taxable Income = 95 | Evade = Yes)**. You should treat Taxable Income as Continuous attribute with the Gaussian Distribution.

Round your answer to the third decimal points.

| *Tid* | Refund | Marital Status | Taxable Income | Evade |
|---|---|---|---|---|
| 1 | Yes | Single | 125K | No |
| 2 | No | Married | 100K | No |
| 3 | No | Single | 70K | No |
| 4 | Yes | Married | 120K | No |
| 5 | No | Divorced | 95K | Yes |
| 6 | No | Married | 60K | No |
| 7 | Yes | Divorced | 220K | No |
| 8 | No | Single | 85K | Yes |
| 9 | No | Married | 75K | No |
| 10 | No | Single | 90K | Yes |

Answer: 0.048 ✔

The correct answer is: 0.048

Question 7

Correct

5.00 points out of 5.00

Given the data in the Table below, and assuming that the attributes follow the Naive Bayes assumption, compute and compare

1. P1 = P ( Evade = Yes | Taxable Income = 95 ) with
2. P2 = P ( Evade = NO | Taxable Income = 95 )

- Report your answer as the ratio of P1 over P2, i.e., P1 / P2
- You should treat Taxable Income as Continuous attribute with the Gaussian Distribution.
- Round your answer to the first decimal point.
- Do NOT use white-space character to allow for auto-grading

| *Tid* | Refund | Marital Status | Taxable Income | Evade |
|---|---|---|---|---|
| 1 | Yes | Single | 125K | No |
| 2 | No | Married | 100K | No |
| 3 | No | Single | 70K | No |
| 4 | Yes | Married | 120K | No |
| 5 | No | Divorced | 95K | Yes |
| 6 | No | Married | 60K | No |
| 7 | Yes | Divorced | 220K | No |
| 8 | No | Single | 85K | Yes |
| 9 | No | Married | 75K | No |
| 10 | No | Single | 90K | Yes |

Answer: 2.8 ✔

```
vec <- c(95, 85, 90)
mean(vec)
sd (vec)
atr <- 95
const <- 1.0 / (sqrt (2*pi) * sd(vec))
exp.value <- exp((-((atr - mean(vec))^2) / (2 * sd(vec)^2)))
probability <- const * exp.value
probability
class.prob <- 3.0 / 10
class.pred1 <- probability * class.prob

vec <- c(125, 100, 70, 120, 60, 220, 75)
mean(vec)
sd (vec)
atr <- 95
const <- 1.0 / (sqrt (2*pi) * sd(vec))
exp.value <- exp((-((atr - mean(vec))^2) / (2 * sd(vec)^2)))
probability <- const * exp.value
probability
class.prob <- 7.0 / 10
class.pred2 <- probability * class.prob
class.pred1 / class.pred2
```

The correct answer is: 2.8

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

Jump to...

(DUE: 02/06/2020): SUBMIT: HW: Logistic Regression ►