

---

## DSC 40A - Discussion 07 - Naive Bayes and Conditional Independence

---

March 10, 2020

---

### Problem 1.

In parts of the world other than San Diego, the weather changes from day to day. In these places, people try to guess tomorrow's weather using the current conditions.

Weather data for 20 random days in Columbus, Ohio are recorded below, along with the next day's weather (rainy, cloudy, or sunny).

Suppose that today's humidity is  $> 50\%$ , the temperature is hot, and the air pressure is low. Use naïve Bayes to predict whether tomorrow will be rainy, cloudy, or sunny. Show your work.

| Next Day's Weather | Humidity | Temperature | Air Pressure |
|--------------------|----------|-------------|--------------|
| Rainy              | $> 50\%$ | Cool        | Low          |
| Rainy              | $> 50\%$ | Hot         | Low          |
| Rainy              | $> 50\%$ | Cool        | Low          |
| Rainy              | 25%-50%  | Hot         | High         |
| Rainy              | 25%-50%  | Hot         | Low          |
| Rainy              | 25%-50%  | Cool        | Low          |
| Rainy              | 25%-50%  | Cool        | Low          |
| Rainy              | $< 25\%$ | Cool        | Low          |
| Rainy              | $< 25\%$ | Hot         | Low          |
| Rainy              | $< 25\%$ | Hot         | High         |
| Cloudy             | $> 50\%$ | Cool        | Low          |
| Cloudy             | $> 50\%$ | Cool        | Low          |
| Cloudy             | 25%-50%  | Hot         | High         |
| Cloudy             | $< 25\%$ | Cool        | High         |
| Cloudy             | $< 25\%$ | Cool        | Low          |
| Sunny              | $> 50\%$ | Cool        | Low          |
| Sunny              | $> 50\%$ | Hot         | High         |
| Sunny              | $> 50\%$ | Cool        | High         |
| Sunny              | 25%-50%  | Hot         | High         |
| Sunny              | $< 25\%$ | Hot         | High         |

Prediction: tomorrow will be      ☐ rainy      ☐ cloudy      ☐ sunny

### Problem 2.

This problem will illustrate that independence and conditional independence are not related, in the sense that neither property implies the other.

Consider the sample space  $S = \{1, 2, 3, 4, 5, 6\}$  with associated probability distribution  $P(s) = \frac{1}{6}$  for each  $s$  in  $S$ . You can think of  $S$  as representing the possible outcomes of rolling a single die.

For each part below, define events  $A, B, C$  in this sample space that satisfy the given requirements. Demonstrate that the requirements are satisfied by computing the appropriate probabilities. Make sure to choose events that are neither impossible nor certain, that is,  $0 < P(A), P(B), P(C) < 1$ .

- a)  $A$  and  $B$  are independent.  
 $A$  and  $B$  are conditionally independent given  $C$ .

**Problem 3.**

A box contains two coins: a regular coin and one fake two-headed coin ( $P(H) = 1$ ). I choose a coin at random and toss it twice. Define the following events.

1.  $A$  = First coin toss results in a heads (H).
2.  $B$  = Second coin toss results in a heads (H).
3.  $C$  = Coin 1 (regular) has been selected.

Find  $P(A|C)$ ,  $P(B|C)$ ,  $P(A \cap B|C)$ ,  $P(A)$ ,  $P(B)$ ,  $P(A \cap B)$ . Show that  $A$  and  $B$  are NOT independent, but they are conditionally independent given  $C$ .