# Randomness, Bayes

***Submit a PDF of your answers to Canvas***

1. ***Laws of Probability.*** Prove or disprove if the following statements are true in general:

   **a)** $\mathbb{P}(A, B^c) = \mathbb{P}(A) - \mathbb{P}(A, B) + \mathbb{P}(A^c) - 1$

   **b)** $\mathbb{P}(A) = 1 - \mathbb{P}(A^c, B) - \mathbb{P}(A^c, B^c)$

   **c)** $\mathbb{P}(A^c \cup B^c) \leq 2.1 - (\mathbb{P}(A) + \mathbb{P}(B))$

   1. Not true, because P(A)+P(A^c)=1,
   which yields P(A,B)+P(A,B^c)=0, which is not necessarily the case
   2. True. P(A^c,B)+P(A^c,B^c)=P(A^c)=1-P(A)
   3. True. By de Morgan's rule, P(A^c u B^c) = 1-P(A,B)
   P(A)+P(B)-P(A,B)=P(A u B) <= 1 < 1.1 as desired

2. ***Bayes Rule and Binary Classifiers.*** There's a new global pandemic, and your roommate was chosen at random to try a new ML based binary classifier that predicts if they have an infectious viral disease.

   The classifier takes in information about symptoms: current temperature (a real number), and a handful of 0/1 variables: cough, shortness of breath, fatigue, muscle/body aches, headache, loss of taste, sore throat, congestion, persistent pain in your chest, blushed lips, and swollen toes.

   The classifier predicts that your roommate is infected. You learn some statistics about the classifier: if infected, the probability of a positive prediction is 0.95, and if not infected, the probability of a negative prediction is 0.99. The current rate of infection in your population is somewhat rare, afflicting only 1 in 1000 people.

   **a)** What is the probability your roommate has the infectious disease (without further information)? Should your roommate be concerned? P(Infected | Positive)=0.95/(0.95+9.99)=95/1094

   **b)** Recall that the *risk* is defined as the average value of the *loss*. If the loss is the misclassification loss (which assigns a 0 or 1 if classification was correct or incorrect), what is the risk of the classifier? 10.04/1000 (10.04 in 1000 is misclassified on average)

   **c)** What is the risk of a classifier that just assigns everyone as *not* having the infection?

   1/1000 (since only 1 in 1000 is actually infected)

Confusion Matrix

| | Positive | Negative | Total |
|---|---|---|---|
| Infected | 0.95 | 0.05 | 1 |
| Not Infected | 9.99 | 989.01 | 999 |