Name: [redacted] **GSI: Mcphail-Snyder** **DISC #: 22458**

Math 55, Handout 17.

THE PROBABILISTIC METHOD.

1.1. If the probability that an element chosen at random from a set S **does not** have a particular property is less than 1, then

The probability that it does have this property is more than 0. Hence there exists elements with this property in S.

Q1. Does the underlying probability distribution have to be uniform for the probabilistic method to work? Explain your answer.

Yes. As it is stated, the element is chosen at *random*, and so the elements in the sample space are assigned a uniform distribution so that $p(\overline{E}) < 1 \Leftrightarrow p(E) > 0 \Leftrightarrow \frac{|E|}{|S|} > 0 \Leftrightarrow |E| > 0$, which proves that E is nonempty and so that there exists elements in S that have the property.

1.2. [Theorem.] If $k \in \mathbb{N} \setminus \{1\}$, then $R(k,k) \geq \underline{2^{k/2}}$.

Q2. Sketch the main steps of the proof given for this Theorem.

- What probability distribution is used?

  The uniform distribution.

- What are the events E_i? How many of them are there?

  The event E_i is that all k people in $S_i = \{\text{the } i\text{th set of } k \text{ elements}\}$ are either mutual friends or mutual enemies.

- What is the probability of each E_i?

  $p(E_i) = 2(\frac{1}{2})^{\frac{k(k-1)}{2}}$, since the probability that all k people in S_i are friends is the same as the probability that all k people are enemies; so the probability is twice these two which is 2 times $\frac{1}{2}$ multiplied by itself the number of pairs that are in S_i since the probability that each edge is either friend or enemy is $\frac{1}{2}$ and each edge is probabalistically independent from every other; so the proability that they are all of the same type is such a product of the same number.

- What is the initial upper bound on $p(\cup_i E_i)$?

  Each E_i has the same probability and so $p(\cup_i E_i) \leq \sum_i p(E_i) = \binom{n}{k} 2(\frac{1}{2})^{\frac{k(k-1)}{2}}$ where n is the number of people of the party S and so there are $\binom{n}{k}$ k-element subsets of S, giving that the sum is $\binom{n}{k}$ times the summand.

- Why is that upper bound strictly less than 1?

  For $k \geq 4$, we use the fact that in general $\binom{n}{k} \leq n^k/2^{k-1}$, so

  $$\binom{n}{k} 2(\frac{1}{2})^{\frac{k(k-1)}{2}} \leq \frac{n^k}{2^{k-1}} 2(\frac{1}{2})^{k(k-1)/2}$$

  now the bound we are looking for occurs when $n < 2^{k/2}$, so if this holds then,
  $\frac{n^k}{2^{k-1}} 2(\frac{1}{2})^{k(k-1)/2} < \frac{2^{k^2/2}}{2^{k-1}} 2(\frac{1}{2})^{k(k-1)/2} = 2^{k^2/2-k+1+1-k(k-1)/2} = 2^{k^2/2-k+1+1-k^2/2+k/2} = 2^{-k+1+1+k/2} = 2^{2-k/2}$
  which is less than 1 since here $k \geq 4$.

- So what if it is strictly less than 1?

  Since the probability that for any given set of $n < 2^{k/2}$ people, there exists a k element subset that consists of mutual friends or mutual enemies is less than 1, it follows that the probability that there exists a set of $n < 2^{k/2}$ people in which every k element subset is not either only of purely mutual friends or mutual enemies is more than 0, hence since the distribution is uniform, the number of sets of people that have k element subsets that are not purely one type is positive, and so there exist such sets of people, as long as the number of people in the set is less than $2^{k/2}$.

- What is the punchline?

  Any set of $n < 2^{k/2}$ people has at least one k element subset that is not of purely mutual enemies or mutual friends.

- What did you learn from this proof?

  That proving that a set E is nonempty can be nontrivially done with probability theory.

BAYES' THEOREM.

2.1. [Bayes' Theorem.] Suppose E and F are events from a sample space S such that $\underline{p(E) \neq 0 \text{ and } p(F) \neq 0}$. Then

$$p(F|E) = \frac{p(F|E)p(F)}{p(E)} = \frac{p(F|E)p(F)}{p(E \cap F) + p(E \cap \overline{F})} = \frac{p(F|E)p(F)}{p(F)p(E|F) + p(\overline{F})p(E|\overline{F})}$$

Q3. You are handed two boxes. You cannot see what is inside but you have reliable information that one of the boxes contains two green and ten red balls and the other contains eight green and four red balls. You select a ball from one of the boxes at random and take it out. It is red. Which of the two boxes is it more likely to have come from? What is that probability?

It is more likeley to have come from the 2 green + 10 red ball box. Call the event of picking from this box A and call the event of picking from the other box B. Then the probability that one picks a red ball from box A is $p(r \cap A) = p(r|A)p(A) = \frac{10}{12} * \frac{1}{2} = \frac{5}{12}$, wher r is the event that one picks a red ball. While $p(r \cap B) = p(r|B)p(B) = \frac{4}{12} * \frac{1}{2} = \frac{1}{6}$, which is less than the latter probability and so it is more likely that a red ball would be chosen from box A. From I could tell, this was what the problem was asking for, not for $p(r)$.

2.1. [Generalized Bayes' Theorem.] Suppose that $E \subseteq S$ and $F_1, \ldots, F_n$ are mutually exclusive events from a sample space S such that $p(E_i) \neq 0$ and $p(F_i) \neq 0$ for all $i = 1, 2, \ldots n$. Then

$$p(F_j|E) = \frac{p(E|F_j)p(F_j)}{\sum_{i=1}^{n} p(E|F_i)p(F_i)}$$

Q4. Suppose that 8% of the patients tested in a particular clinic are infected with HIV. Furthermore, suppose that when a blood test for HIV is given, 98% of the patients infected with HIV test positive and that 3% of the patients not infected with HIV test positive. What is the probability that

(a) a patient testing positive for HIV with this test is infected with it?

$$p(H|+) = \frac{p(H)p(+|H)}{p(+)} = \frac{p(H)p(+|H)}{p(H)p(+|H) + p(\overline{H})p(+|\overline{H})} = \frac{0.08 * 0.98}{0.08 * 0.98 + (1-0.08)(0.03)} = 0.7396$$

(b) a patient testing positive for HIV with this test is not infected by it?

$$p(\overline{H}|+) = \frac{p(\overline{H})p(+|\overline{H})}{p(+)} = \frac{p(\overline{H})p(+|\overline{H})}{p(\overline{H})p(+|\overline{H}) + p(H)p(+|H)} = \frac{(1-0.08) * 0.03}{(1-0.08) * 0.03 + (0.08)(0.98)} = 0.260 = 1 - p(H|+)$$

(c) a patient testing negative for HIV with this test is infected with it?

$$p(H|\overline{+}) = \frac{p(\overline{+} \cap H)}{p(\overline{+})} = \frac{1 - p(+ \cup \overline{H})}{p(\overline{+})} = \frac{1 - (p(+) + p(\overline{H}) - p(+ \cap \overline{H}))}{1 - p(+)} = 1 - \frac{p(\overline{H}) - p(+ \cap \overline{H}))}{1 - p(+)}$$

$$= 1 - \frac{p(\overline{H}) - p(\overline{H})p(+|\overline{H}))}{1 - p(+)} = 1 - \frac{(1-0.08)(1-0.03))}{1 - ((1-0.08) * 0.03 + (0.08)(0.98))} = 0.0018$$

(d) a patient testing negative for HIV with this test is not infected with it?

$$p(\overline{H}|\overline{+}) = 1 - p(H|\overline{+}) = 0.998$$