Probability Definitions _____

(from Appendix C.3)

• A *random variable* is a variable that takes on one of several values, each with some probability. (Example: if *X* is the outcome of the role of a die, *X* is a random variable)

• The *expected value* of a random variable, *X* is defined as:

$$E(X) = \sum_{x} x * P(X = x)$$

(Example if X is the outcome of the role of a three sided die,

$$E(X) = 1 * (1/3) + 2 * (1/3) + 3 * (1/3)$$

= 2

Probability Definitions _____

- Two events *A* and *B* are *mutually exclusive* if *A*∩*B* is the empty set (Example: *A* is the event that the outcome of a die is 1 and *B* is the event that the outcome of a die is 2)
- Two random variables X and Y are *independent* if for all x and y, P(X = x and Y = y) = P(X = x)P(Y = y) (Example: let X be the outcome of the first role of a die, and Y be the outcome of the second role of the die. Then X and Y are independent.)

CS 361, Lecture 11

Jared Saia University of New Mexico

Outline _____

- Birthday Paradox
- Analysis of Randomized Quicksort

2

Probability Definitions _____

___ Example ____

- An *Indicator Random Variable* associated with event *A* is defined as:
 - -I(A) = 1 if A occurs
 - -I(A) = 0 if A does not occur
- Example: Let A be the event that the role of a die comes up 2. Then I(A) is 1 if the die comes up 2 and 0 otherwise.

- For $1 \leq i \leq n$, let X_i be the outcome of the *i*-th role of three-sided die
- Then

$$E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = 2n$$



- Let X and Y be two random variables
- Then E(X + Y) = E(X) + E(Y)
- (Holds even if X and Y are not independent.)
- More generally, let X_1, X_2, \ldots, X_n be *n* random variables
- Then

$$E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$$

- Indicator Random Variables and Linearity of Expectation used together are a very powerful tool
- The "Birthday Paradox" illustrates this point
- To analyze the run time of quicksort, we will also use indicator r.v.'s and linearity of expectation (analysis will be similar to "birthday paradox" problem)

"Birthday Paradox"

__ Analysis ____

- Assume there are k people in a room, and n days in a year
- Assume that each of these k people is born on a day chosen uniformly at random from the n days
- Q: What is the expected number of pairs of individuals that have the same birthday?
- We can use indicator random variables and linearity of expectation to compute this

- Let *X* be a random variable giving the number of pairs of people with the same birthday
- We want E(X)

• Then
$$X = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} X_{i,j}$$

• So $E(X) = E(\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} X_{i,j})$



- For all $1 \le i < j \le k$, let $X_{i,j}$ be an indicator random variable defined such that:
 - $\ X_{i,j} = 1$ if person i and person j have the same birthday $\ X_{i,j} = 0$ otherwise
- Note that for all *i*, *j*,

$$E(X_{i,j}) = P(\text{person i and j have same birthday})$$

= $1/n$

$$E(X) = E(\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} X_{i,j})$$

= $\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} E(X_{i,j})$
= $\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} 1/n$
= $\binom{k}{2} 1/n$
= $\frac{k(k-1)}{2n}$

The second step follows by Linearity of Expectation

Reality Check _____

In-Class Exercise _____

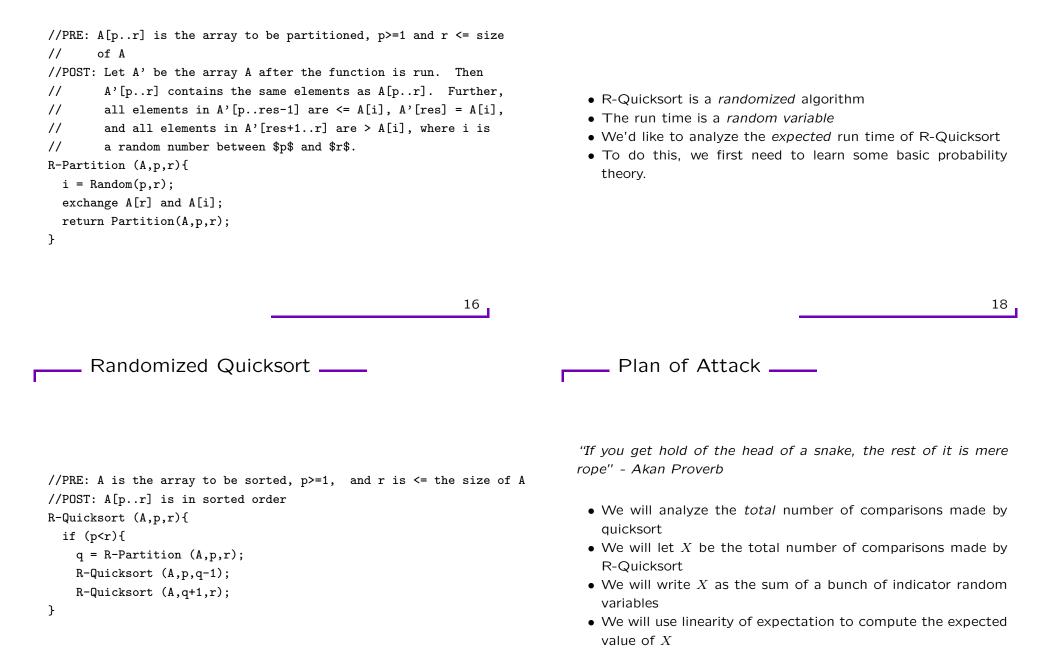
- Thus, if $k(k-1) \ge 2n$, expected number of pairs of people with same birthday is at least 1
- Thus if have at least $\sqrt{2n} + 1$ people in the room, can expect to have at least two with same birthday
- For *n* = 365, if *k* = 28, expected number of pairs with same birthday is 1.04

- Q1: Write the expected value of X as a function of the X_{i,j,k} (use linearity of expectation)
- Q2: What is $E(X_{i,j,k})$?
- Q3: What is the total number of groups of three people out of *k*?
- Q4: What is E(X)?



- Assume there are k people in a room, and n days in a year
- Assume that each of these k people is born on a day chosen uniformly at random from the n days
- Let X be the number of groups of *three* people who all have the same birthday. What is E(X)?
- Let $X_{i,j,k}$ be an indicator r.v. which is 1 if people i,j, and k have the same birthday and 0 otherwise

```
___ Analysis ____
```



Notation ____

• Let A be the array to be sorted

- Let z_i be the *i*-th smallest element in the array A
- Let $Z_{i,j} = \{z_i, z_{i+1}, \dots, z_j\}$

• Q1: So what is $E(X_{i,j})$?

• A1: It is $P(z_i \text{ is compared to } z_j)$

Questions _____

- Q2: What is $P(z_i \text{ is compared to } z_j)$?
- A2: It is:

 $P(\text{either } z_i \text{ or } z_j \text{ are the first elems in } Z_{i,j} \text{ chosen as pivots})$

- Why?
 - If no element in $Z_{i,j}$ has been chosen yet, no two elements in $Z_{i,j}$ have yet been compared, and all of $Z_{i,j}$ is in same list
 - If some element in $Z_{i,j}$ other than z_i or z_j is chosen first, z_i and z_j will be split into separate lists (and hence will never be compared)

20

Indicator Random Variables _____

- Let $X_{i,j}$ be 1 if z_i is compared with z_j and 0 otherwise
- Note that $X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{i,j}$
- Further note that

$$E(X) = E(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{i,j}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{i,j})$$

— More Questions ——

• Q: What is

 $P(\text{either } z_i \text{ or } z_j \text{ are first elems in } Z_{i,j} \text{ chosen as pivots})$

- A: $P(z_i \text{ chosen as first elem in } Z_{i,j}) + P(z_i \text{ chosen as first elem in } Z_{i,j})$
- Further note that number of elems in $Z_{i,j}$ is j i + 1, so

$$P(z_i \text{ chosen as first elem in } Z_{i,j}) = rac{1}{j-i+1}$$

and

$$P(z_j \text{ chosen as first elem in } Z_{i,j}) = \frac{1}{j-i+1}$$

• Hence

$$P(z_i \text{ or } z_j \text{ are first elems in } Z_{i,j} \text{ chosen as pivots}) = \frac{2}{j-i+1}$$

23

22

____ Questions _____

$$E(X_{i,j}) = P(z_i \text{ is compared to } z_j)$$
 (1)

 $= O(n \log n)$

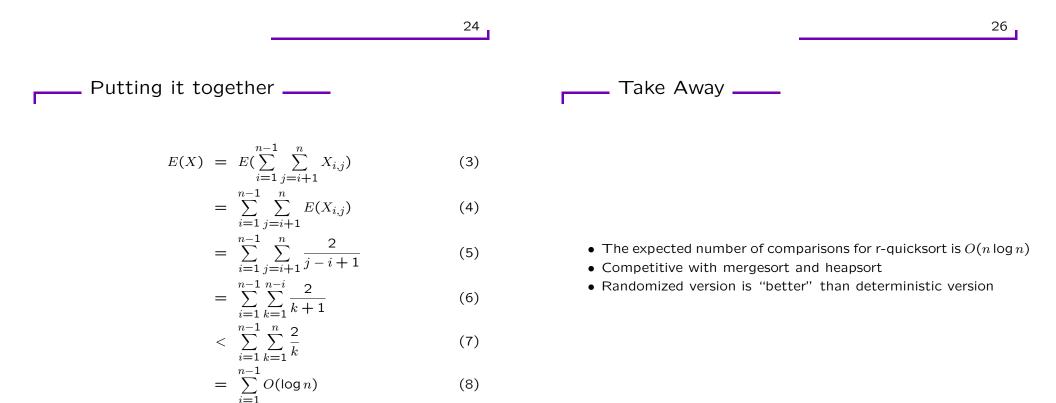
$$= \frac{2}{j-i+1} \tag{2}$$

• Q: Why is
$$\sum_{k=1}^{n} \frac{2}{k} = O(\log n)$$
?
• A:

$$\sum_{k=1}^{n} \frac{2}{k} = 2 \sum_{k=1}^{n} 1/k \qquad (10)$$

$$\leq 2(\ln n + 1) \qquad (11)$$

• Where the last step follows by an integral bound on the sum (p. 1067)



(9) 25 Todo _____

• Finish Chapter 7

28