

# INDICATOR RANDOM VARIABLES

(IRV)

(taking value 0 or 1)

We already saw this:  $Y$ : parity of rolling one die.

Another example: flip a coin 10 times.

$X$  = #times we see pattern HT

$$E[X] = ?$$

# INDICATOR RANDOM VARIABLES

flip a coin 10 times.

$X = \#$  times we see pattern HT

HT could appear at flips 1&2, or 2&3, ..., or 9 & 10

Define r.v.  $X_i = \begin{cases} 1 & \text{if flips } i \text{ \& } i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases}$

Notice  $X_1$  &  $X_2$  are not independent.  $P(X_i=1) = \frac{1}{4}$   
 $P(X_1 \wedge X_2) = 0$

$$X = X_1 + X_2 + \dots + X_9$$

$$\begin{aligned} E[X] &= E[X_1 + X_2 + \dots + X_9] \\ &= E[X_1] + E[X_2] + \dots + E[X_9] \end{aligned}$$

linearity of expectation

$$\begin{aligned} E[X_i] &= 0 \cdot P(X_i=0) + 1 \cdot P(X_i=1) = \frac{1}{4} \\ &= 9 \cdot \frac{1}{4} \end{aligned}$$

# INDICATOR RANDOM VARIABLES

The hat-check problem (a.k.a. coat-check)

- ◆  $n$  people at a party leave their hats with an attendant
- ◆ The attendant gives hats back randomly.

How many people do we expect to get their own hats back?

# The hat-check problem

$$X = \# \text{ people who get their own hat back} = \sum_{k=1}^n X_k$$

$$X_k = \begin{cases} 1 & \text{if person } k \text{ gets their own hat back} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = E\left[\sum_{k=1}^n X_k\right]$$

$$= \sum_{k=1}^n E[X_k]$$

$$= \sum_{k=1}^n \frac{1}{n} = \mathbf{1}$$

linearity of expectation

$$E[X_k] = \frac{1}{n} \quad \begin{array}{l} \text{(random)} \\ P(X_k = 1) \end{array}$$

The hiring problem: you need one assistant.

◆  $n$  candidates, interviewed in random order.

◆ No 2 equally skilled.

◆ any time you interview someone better than all previous, you hire the new person & fire the current assistant.

How many people do you expect to hire?

# The hiring problem

$$X = \# \text{ people you will hire} = \sum_{k=1}^n X_k$$

$$X_k = \begin{cases} 1 & \text{if you hire candidate } k \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = E\left[\sum_{k=1}^n X_k\right]$$

$$= \sum_{k=1}^n E[X_k]$$

$$= \sum_{k=1}^n \frac{1}{k} \leq 1 + \ln n$$

linearity of expectation

$$E[X_k] = \frac{1}{k}$$

person  $k$  is hired  
iff better than  
all  $k-1$  previous

## The birthday problem

How many people do we need in a room so that we expect to have (at least) one birthday match?

# The birthday problem

$X$  = # birthday matches among  $n$  people

For what  $n$  do we get  $E[X] \geq 1$  ?

↳ Set up  $E[X]$  as function of  $n$

# The birthday problem

$X$  = # birthday matches among  $n$  people

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}$$

all  $\binom{n}{2}$  pairs

$$X_{ij} = \begin{cases} 1 & \text{if persons } i \text{ \& } j \text{ match} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X_{ij}] = \frac{1}{365}$$

$$E[X] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^n E[X_{ij}]$$

linearity of expectation  
we said we want  $E[X]=1$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{1}{365} = \frac{n \cdot (n-1)}{2} \cdot \frac{1}{365} = 1 \Rightarrow n \approx 28$$

□

$P(\geq 2 \text{ people in a group of } k \text{ have same birthday})$

(new problem)  
Not about  
expected value or IRV

$= 1 - P(\text{nobody has the same birthday})$

$$= 1 - \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \dots \cdot \frac{365 - (k-1)}{365} = 1 - \frac{365!}{(365-k)! 365^k}$$

Diagram illustrating the probability calculation for the birthday problem. The expression is written in red ink. The first term is  $\frac{365}{365}$ , the second is  $\frac{364}{365}$ , the third is  $\frac{363}{365}$ , followed by an ellipsis, and the final term is  $\frac{365 - (k-1)}{365}$ . The entire product is subtracted from 1. The final simplified form is  $1 - \frac{365!}{(365-k)! 365^k}$ . Green curved lines above the fractions connect them to labels: "person #1" above the first fraction, "person #2" above the second, "person #3" above the third, and "person #k" above the last fraction.

$P(\geq 2 \text{ people in a group of } k \text{ have same birthday})$

$$k=2 \rightarrow P \sim 0.27\% \left(\frac{1}{365}\right)$$

$$k=4 \rightarrow P \sim 1.64\%$$

$$k=23 \rightarrow P \sim 50.73\%$$

$$k=30 \rightarrow P \sim 70.6\%$$

$$k=70 \rightarrow P \sim 99.9\%$$

$$k \sim 116 \rightarrow P \sim 1 - \frac{1}{10^9}$$

$$k=300 \rightarrow P \sim 1 - \frac{1}{10^{80}}$$

$(10^{80} \sim \# \text{ atoms in universe})$

$$(k > 365 \rightarrow P=1)$$