INDICATOR RANDOM VARIABLES

(taking value 0 or 1)

INDICATOR RANDOM VARIABLES

(taking value 0 or 1)

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

INDICATOR RANDOM VARIABLES

(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

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

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 = ?

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

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

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

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

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

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

$P(X_1 \wedge X_2) = 0$

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. $P(X_i=1)=\frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(X) = E(X_1 + X_2 + \cdots + X_9)$$

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. $P(X_i = 1) = \frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(x) = E(X_1 + X_2 + \cdots + X_9)$$

linearity of expectation

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. $P(X_i = 1) = \frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(X) = E(X_1 + X_2 + \cdots + X_9)$$

$$= E(X_1) + E(X_2) + \cdots E(X_9)$$

linearity of expectation

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. $P(X_i=1)=\frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(X) = E(X_1 + X_2 + \cdots + X_9)$$

$$= E(X_1) + E(X_2) + \cdots E(X_9)$$

linearity of expectation

$E(X_i) =$?

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. $P(X_i = 1) = \frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(X) = E(X_1 + X_2 + \cdots + X_9)$$

linearity of expectation

$$= E(X_1) + E(X_2) + \cdots E(X_9)$$

$$E(X_i) = 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1)$$

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. $P(X_i = 1) = \frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(X) = E(X_1 + X_2 + \cdots + X_9)$$

$$= E(X_1) + E(X_2) + \cdots E(X_9)$$

linearity of expectation

$$E(X_i) = \cancel{0 \cdot P(X_i = 0) + 1} \cdot \underline{P(X_i = 1)}$$

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. $P(X_i = 1) = \frac{1}{4}$

$P(X_1 \wedge X_2) = 0$

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

$$E(X) = E(X_1 + X_2 + \cdots + X_9)$$

$$= E(X_1) + E(X_2) + \cdots E(X_9)$$

linearity of expectation

$$E(X_i) = \cancel{0 \cdot P(X_i = 0) + 1} \cdot \underline{P(X_i = 1)} = \frac{1}{4}$$

$$= 9 \cdot \frac{1}{4}$$

INDICATOR RANDOM VARIABLES

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

INDICATOR RANDOM VARIABLES

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

- n people leave their hats with an attendant, & get a ticket = number for retrieval.

INDICATOR RANDOM VARIABLES

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

- n people leave their hats with an attendant, & get a ticket = number for retrieval.

- The attendant loses all ticket info & gives hats back randomly.

INDICATOR RANDOM VARIABLES

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

- n people leave their hats with an attendant,
  & get a ticket = number for retrieval.

- The attendant loses all ticket info
  & gives hats back randomly.

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

INDICATOR RANDOM VARIABLES

X = # people who get their own hat back

INDICATOR RANDOM VARIABLES

X = # people who get their own hat back

$X_k = ?$

INDICATOR RANDOM VARIABLES

X = # people who get their own hat back

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

INDICATOR RANDOM VARIABLES

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

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

INDICATOR RANDOM VARIABLES

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

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

$$E(X_k) = ?$$

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{n}$ (random)

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{n}$ (random)

$E(X) = \ ?$

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{n}$ (random)

$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = ?$

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{n}$ (random)

$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$ linearity of expectation

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{n}$ (random)

$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$ linearity of expectation

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

INDICATOR RANDOM VARIABLES

The hiring problem: you need one assistant.

INDICATOR RANDOM VARIABLES

The hiring problem: you need one assistant.

- n candidates, interviewed in <u>random</u> order.

INDICATOR RANDOM VARIABLES

The hiring problem: you need one assistant.

- n candidates, interviewed in <u>random</u> order.
- No 2 equally skilled.

INDICATOR RANDOM VARIABLES

The hiring problem: you need one assistant.

- n candidates, interviewed in <u>random</u> order.
- No 2 equally skilled.
- any time you interview someone better than all previous, you hire the new person & fire the current assistant.

INDICATOR RANDOM VARIABLES

The hiring problem: you need one assistant.

- n candidates, interviewed in <u>random</u> 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?

INDICATOR RANDOM VARIABLES

X = # people you expect to hire

INDICATOR RANDOM VARIABLES

X = # people you expect to hire

$X_k = ?$

INDICATOR RANDOM VARIABLES

X = # people you expect to hire = ?

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

INDICATOR RANDOM VARIABLES

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

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

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = ?$

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{k}$ (k is hired <u>iff</u> better than all $k-1$ previous)

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{k}$ (k is hired iff better than all k-1 previous)

$E(X) = \,?$

INDICATOR RANDOM VARIABLES

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

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

$$E(X_k) = \frac{1}{k}$$ (k is hired <u>iff</u> better than all k-1 previous)

$$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = ?$$

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{k}$ (k is hired <u>iff</u> better than all $k-1$ previous)

$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$ linearity of expectation

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{k}$ (k is hired <u>iff</u> better than all k-1 previous)

$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$ linearity of expectation

$= \sum_{k=1}^{n} \frac{1}{k}$ ...?

INDICATOR RANDOM VARIABLES

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

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

$E(X_k) = \frac{1}{k}$ (k is hired <u>iff</u> better than all $k-1$ previous)

$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$ linearity of expectation

$$= \sum_{k=1}^{n} \frac{1}{k} = \ln n + O(1) < \ln n + 1$$

INDICATOR RANDOM VARIABLES

The birthday problem (new variant)

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

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people

What should our I.R.V. be ? $X_?$

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people

$X_{ij} = \ ?$

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people

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

INDICATOR RANDOM VARIABLES

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

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

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \ ?$$

INDICATOR RANDOM VARIABLES

$X = \#$ birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = \ ?$$

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right) = \ ?$$

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}}_{\text{all } \binom{n}{2} \text{ pairs}}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij})$$ linearity of expectation

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij}) \quad \text{linearity of expectation}$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{365}$$

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij})$$ linearity of expectation

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

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij})$$ linearity of expectation

we said we want $E[X]=1$

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

INDICATOR RANDOM VARIABLES

X = # birthday matches among n people $= \underbrace{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n}}_{\text{all } \binom{n}{2} \text{ pairs}} X_{ij}$

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

$$E(X_{ij}) = \frac{1}{365}$$

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij})$$ linearity of expectation

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

we said we want $E[X]=1$

NOT 23