If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?

Hmm...

\[ \sum_k k \Pr(k \text{ letters end up in correct envelopes}) = \sum_k k \ldots \text{aargh!} \ldots \]

On average, in class of size \( m \), how many pairs of people will have the same birthday?

\[ \sum_k k \Pr(\text{exactly } k \text{ collisions}) = \sum_k k \ldots \text{aargh!!!!} \ldots \]
The new tool is called “Linearity of Expectation”

To use this new tool, we will also need to understand the concept of a Random Variable

Today’s lecture: not too much material, but need to understand it well

Random Variable

Let S be a sample space in a probability distribution.
A Random Variable is a real-valued function on S.

Examples:

- \( X = \) value of white die in a two-dice roll
  \( X(3,4) = 3, \quad X(1,6) = 1 \)

- \( Y = \) sum of values of the two dice
  \( Y(3,4) = 7, \quad Y(1,6) = 7 \)

- \( W = (\text{value of white die})^{\text{value of black die}} \)
  \( W(3,4) = 3^4, \quad W(1,6) = 1^6 \)

Tossing a Fair Coin n Times

S = all sequences of \( \{H, T\}^n \)
D = uniform distribution on S
\( \Rightarrow D(x) = (\frac{1}{2})^n \quad \text{for all} \quad x \in S \)

Random Variables (say n = 10)

- \( X = \) # of heads
  \( X(\text{HHHTTHTHTT}) = 5 \)

- \( Y = \) (1 if #heads = #tails, 0 otherwise)
  \( Y(\text{HHHTTHTHTT}) = 1, \quad Y(\text{THHHHHTTTTT}) = 0 \)
Notational Conventions

Use letters like A, B, E for events
Use letters like X, Y, f, g for R.V.’s
R.V. = random variable

Two Views of Random Variables

Think of a R.V. as
A function from S to the reals $\mathbb{R}$
Or think of the induced distribution on $\mathbb{R}$
Randomness is “pushed” to the values of the function

Two Coins Tossed

$X: \{TT, TH, HT, HH\} \rightarrow \{0, 1, 2\}$ counts the number of heads

Distribution on the reals

It’s a Floor Wax And a Dessert Topping

It’s a function on the sample space $S$
It’s a variable with a probability distribution on its values
You should be comfortable with both views
From Random Variables to Events

For any random variable $X$ and value $a$, we can define the event $A$ that $X = a$

$$Pr(A) = Pr(X=a) = Pr\{x \in S | X(x) = a\}$$

Two Coins Tossed

$X: \{TT, TH, HT, HH\} \rightarrow \{0, 1, 2\}$ counts # of heads

\[
\begin{array}{c}
\text{TT} & \frac{1}{4} \\
\text{TH} & \frac{1}{4} \\
\text{HT} & \frac{1}{4} \\
\text{HH} & \frac{1}{4}
\end{array}
\]

Pr($X = 1$) = Pr($\{x \in S | X(x) = 1\}$) = \frac{1}{2}

From Events to Random Variables

For any event $A$, can define the indicator random variable for $A$:

$$X_A(x) = \begin{cases} 
1 & \text{if } x \in A \\
0 & \text{if } x \notin A
\end{cases}$$

Definition: Expectation

The expectation, or expected value of a random variable $X$ is written as $E[X]$, and is

$$E[X] = \sum_{x \in S} Pr(x) X(x) = \sum_{k} k Pr[X = k]$$

$X$ is a function on the sample space $S$ and $X$ has a distribution on its values.
A Quick Calculation…

What if I flip a coin 3 times? What is the expected number of heads?
What are total number of possibilities for 3 coins?

\[ E[X] = \]

But \( \Pr[ X = 1.5 ] = \)

Moral: don’t always expect the expected.
\( \Pr[ X = E[X] ] \)

Type Checking

A Random Variable is the type of thing you might want to know an expected value of

If you are computing an expectation, the thing whose expectation you are computing is a random variable

Indicator R.V.s: \( E[X_A] = \Pr(A) \)

For any event \( A \), can define the indicator random variable for \( A \):

\[
X_A(x) = \begin{cases} 
1 & \text{if } x \in A \\
0 & \text{if } x \notin A 
\end{cases}
\]

\[ E[X_A] = 1 \times \Pr(X_A = 1) = \Pr(A) \]

Adding Random Variables

If \( X \) and \( Y \) are random variables (on the same set \( S \)), then \( Z = X + Y \) is also a random variable

\[ Z(x) = X(x) + Y(x) \]

E.g., rolling two dice.
\( X = 1\)st die, \( Y = 2\)nd die,
\( Z = \) sum of two dice
Adding Random Variables

Example: Consider picking a random person in the world. Let $X =$ length of the person’s left arm in inches. $Y =$ length of the person’s right arm in inches. Let $Z = X + Y$. $Z$ measures the combined arm lengths.

Independence

Two random variables $X$ and $Y$ are independent if for every $a, b$, the events $X = a$ and $Y = b$ are independent.

How about the case of $X =$ 1st die, $Y =$ 2nd die? $X =$ left arm, $Y =$ right arm?

Linearity of Expectation

If $Z = X + Y$, then

$$E[Z] = E[X] + E[Y]$$

Even if $X$ and $Y$ are not independent,

$$E[Z] = \sum_{x \in S} \Pr[x] \cdot Z(x)$$

$$= \sum_{x \in S} \Pr[x] \cdot (X(x) + Y(x))$$

$$= \sum_{x \in S} \Pr[x] \cdot X(x) + \sum_{x \in S} \Pr[x] \cdot Y(x)$$
Linearity of Expectation

E.g., 2 fair flips:
- X = 1st coin # heads, Y = 2nd coin # heads
- Z = X + Y = total # heads

What is E[X]? E[Y]? E[Z]?

<table>
<thead>
<tr>
<th>Event</th>
<th>Pr(X)</th>
<th>Pr(Y)</th>
<th>Pr(Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>TH</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>HT</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>HH</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
</tbody>
</table>

E[X] = ?

E[X] = \sum_{x \in S} k Pr[X = k]

E[X] = \frac{1}{2}, similarly E[Y] = \frac{1}{2} E[Z] = ?

E[Z] =

E[Z] = E[X] + E[Y] =

Linearity of Expectation

E.g., 2 fair flips:
- X = at least one coin is heads
- Y = both coins are heads, Z = X + Y

Are X and Y independent?

What is E[X]? E[Y]? E[Z]?

<table>
<thead>
<tr>
<th>Event</th>
<th>Pr(X)</th>
<th>Pr(Y)</th>
<th>Pr(Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>TH</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>HT</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
<tr>
<td>HH</td>
<td>1/4</td>
<td>1/4</td>
<td>1/2</td>
</tr>
</tbody>
</table>

E[X] = ?

X = 0 is probability none are heads
There is 1 case no heads =

X = 1 is probability at least one is heads
There are 3 cases =

E[X] = probability of 0 heads * 0 heads +
probability of at least 1 head * 1 head

=
\[ E[Y] = ?, \quad E[Z] = ? \]

Y = both coins are heads
\[ E[Y] = \]

Z is both coins are heads and one coin is heads
\[ E[Z] = \quad E[X] + E[Y] = \]

By Induction
\[ E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n] \]

The expectation of the sum =
The sum of the expectations

It is finally time to show off our probability prowess...

If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?

Hmm...
\[ \sum_k k \Pr(k \text{ letters end up in correct envelopes}) = \sum_k k (\ldots\text{aargh!!}\ldots) \]
Use Linearity of Expectation

Let $A_i$ be the event the $i^{th}$ letter ends up in its correct envelope

Let $X_i$ be the indicator R.V. for $A_i$

\[
X_i = \begin{cases} 
1 & \text{if } A_i \text{ occurs} \\
0 & \text{otherwise}
\end{cases}
\]

Let $Z = X_1 + \ldots + X_{100}$

We are asking for $E[Z]$

$E[X_i] = \Pr(A_i) = 1/100$

So $E[Z] = 1$

So, in expectation, 1 letter will be in the same correct envelope

Pretty neat: it doesn’t depend on how many letters!

Question: were the $X_i$ independent?

No! E.g., think of $n=2$

Use Linearity of Expectation

General approach:

View thing you care about as expected value of some R.V

Write this R.V as sum of simpler R.Vs (typically indicator R.Vs)

Solve for their expectations and add them up!

Example

We flip $n$ coins of bias $p$. What is the expected number of heads?

We could do this by summing

\[
\sum_k k \Pr(X = k) = \sum_k k \binom{n}{k} p^k (1-p)^{n-k}
\]

But now we know a better way!
Let $X =$ number of heads when $n$ independent coins of bias $p$ are flipped.

Break $X$ into $n$ simpler RVs:

$$X_j = \begin{cases} 
1 & \text{if the } j\text{th coin is heads} \\
0 & \text{if the } j\text{th coin is tails}
\end{cases}$$

$$E[X] = \sum_a a \times \Pr(X=a)$$

Linearity of Expectation!

Example: 2 fair flips

$X =$ indicator for 1st coin heads

$Y =$ indicator for 2nd coin heads

$XY =$ indicator for “both are heads”

$E[X] = \frac{1}{2}, \ E[Y] = \frac{1}{2}, \ E[XY] =$

What About Products?

If $Z = XY$, then $E[Z] = E[X] \times E[Y]$?

$X =$ indicator for “1st flip is heads”

$Y =$ indicator for “1st flip is tails”

$E[XY] =$

But It’s True If RVs Are Independent

Proof:

$$E[X] = \sum_a a \times \Pr(X=a)$$

$$E[Y] = \sum_b b \times \Pr(Y=b)$$

$$E[XY] = \sum_c c \times \Pr(XY = c)$$

$$= \sum_c \sum_{a,b:ab=c} c \times \Pr(X=a \cap Y=b)$$

$$= \sum_{a,b} ab \times \Pr(X=a \cap Y=b)$$

$$= \sum_{a,b} ab \times \Pr(X=a) \Pr(Y=b)$$

$$= E[X] \ E[Y]$$

Example: 2 fair flips

$X =$ indicator for 1st coin heads

$Y =$ indicator for 2nd coin heads

$XY =$ indicator for “both are heads”

$E[X] = \frac{1}{2}, \ E[Y] = \frac{1}{2}, \ E[XY] =$

$E[X^2] = E[X]^2?$

In fact, $E[X^2] – E[X]^2$ is called the variance of $X$.
Most of the time, though, power will come from using sums

Mostly because Linearity of Expectations holds even if RVs are not independent

On average, in class of size m, how many pairs of people will have the same birthday?

\[ \sum_k k \Pr(\text{exactly } k \text{ collisions}) = \sum_k k (\ldots \text{aargh!} \ldots) \]

Use linearity of expectation

Suppose we have m people each with a uniformly chosen birthday from 1 to 366

\[ X = \text{number of pairs of people with the same birthday} \]

\[ E[X] = ? \]

\[ X = \text{number of pairs of people with the same birthday} \]

Use \( m(m-1)/2 \) indicator variables, one for each pair of people

\[ X_{jk} = 1 \text{ if person } j \text{ and person } k \text{ have the same birthday; else } 0 \]

\[ E[X_{jk}] = ? \]
\( X = \) number of pairs of people with the same birthday
\( X_{jk} = 1 \) if person \( j \) and person \( k \) have the same birthday; else 0
\[
E[X_{jk}] = \frac{1}{366} \cdot 1 + \left(1 - \frac{1}{366}\right) \cdot 0 = \frac{1}{366}
\]

\[
E[X] = E\left[ \sum_{j \leq k \leq m} X_{jk} \right] = \sum_{j \leq k \leq m} E[X_{jk}]
\]

For \( m = 28 \), \( E[X] = \)

---

**Step Right Up...**

You pick a number \( n \in [1..6] \). You roll 3 dice. If any match \( n \), you win $1. Else you pay me $1. Want to play?

**Hmm... let's see**

---

**Analysis**

\( A_i = \) event that \( i \)-th die matches
\( X_i = \) indicator RV for \( A_i \)

Expected number of dice that match:
\[
E[X_1+X_2+X_3] =
\]

But this is not the same as \( Pr(\text{at least one die matches}) \)