15-251

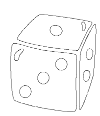
Great Theoretical Ideas in Computer Science

15-251

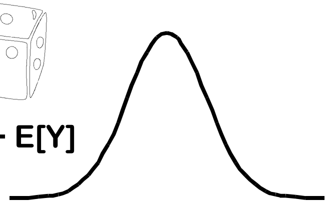
Setting Expectations for Computer Scientists

Probability Theory II

Lecture 12 (September 30, 2010)

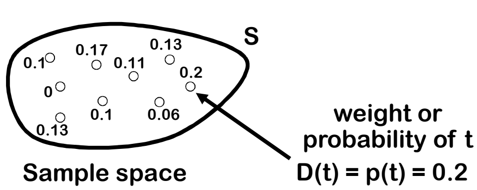


$E[X+Y] = E[X] + E[Y]$



A few things from last time…

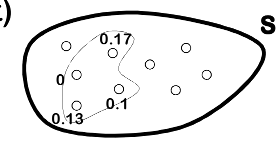
Sample Space



Events

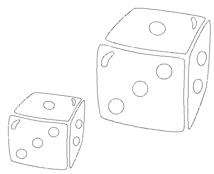
Any set $E \subseteq S$ is called an event

$$Pr_D[E] = \sum_{t \in E} p(t)$$



$$Pr_D[E] = 0.4$$

Binomials and Gaussians…

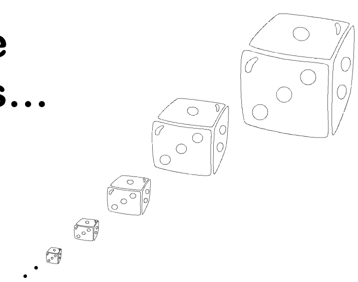


Pr(see k heads in n flips) = $\binom{n}{k} p^k (1-p)^{n-k}$

Given a fair coin (p = ½) $\binom{n}{k} 2^{-n}$

As $n \to \infty$, the plot for Pr(k heads) for p = ½ tends to "bell curve" or "Gaussian/normal distribution"

Infinite sample spaces…



The "Geometric" Distribution

A bias-p coin is tossed until the first time that a head turns up.

sample space S = {H, TH, TTH, TTTH, …}

(shorthand S = {1, 2, 3, 4, …})

$\Pr(k) = (1-p)^{k-1} p$

(sanity check) $\sum_{k\geq 1} \Pr(k) = \sum_{k\geq 1} (1-p)^{k-1} p$

$= p * (1 + (1-p) + (1-p)^2 + \ldots)$

$= p * 1/(1-(1-p)) = 1$

The Geometric Distribution

A bias-p coin is tossed until the first time that a head turns up.

sample space S = {1, 2, 3, 4, … }

E = "first heads at even numbered flip"

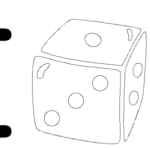
= {2, 4, 6, …} "= {TH, TTTH, TTTTTH, …}"

$\Pr(E) = \sum_{x \text{ in } E} \Pr(x) = \sum_{k \text{ even}} (1-p)^{k-1} p$

$= p(1-p) * (1 + (1-p)^2 + (1-p)^4 + \ldots)$

$= p(1-p) * 1/(1-(1-p)^2) = (1-p)/(2-p)$

Expectations

E[]

"Expectation = (weighted) average"

What is the average height of a 251 student?

Sample space S = all 251 students

(implicit) probability distribution = uniform

average = (sum of heights in S)/|S|

$= (\sum_{t \text{ in } S} \text{Height}(t))/|S|$

$= \sum_{t \text{ in } S} \Pr(t) \text{ Height}(t)$

Expectation

Given a probability distribution D = (S, Pr)

And any function f:S to reals

$$E[\, f \,] = \sum_{t \text{ in } S} f(t) \Pr(t)$$

A function from the sample space to the reals is called a "random variable"

Random Variable

Let S be sample space in a probability distribution

A Random Variable is a function from S to reals

Examples:

X = value of white die in a two-dice roll

X(3,4) = 3, X(1,6) = 1

Y = sum of values of the two dice

Y(3,4) = 7, Y(1,6) = 7

Notational Conventions

Use letters like A, B, C for events

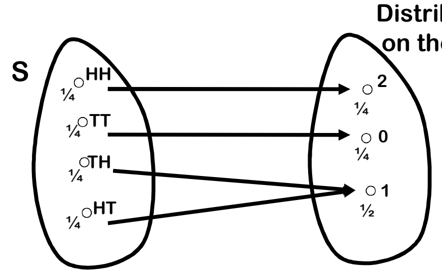
Use letters like X, Y, f, g for R.V.'s

R.V. = random variable

Two Coins Tossed

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

Distribution on the reals!



Two Views of Random Variables

Think of a R.V. as

Input to the function is random

A function from S to the reals R

Or think of the induced distribution on R

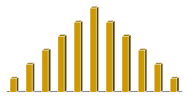
Randomness is "pushed" to the values of the function

Given a distribution, a random variable transforms it into a distribution on reals

Two dice

I throw a white die and a black die.

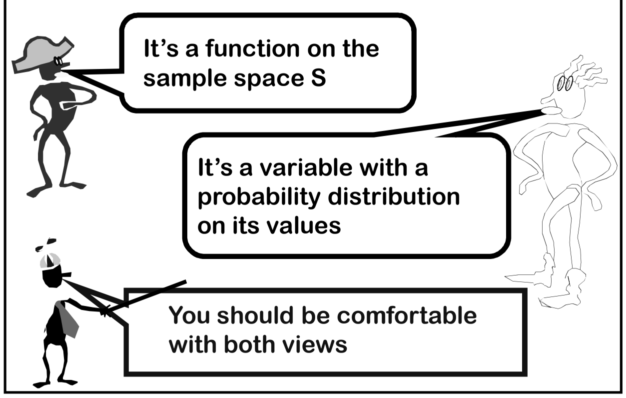
Sample space S =
{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) }



X = sum of both dice

function with X(1,1) = 2, X(1,2) = 3, …, X(6,6)=12

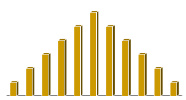
It's a Floor Wax And a Dessert Topping



Two dice

I throw a white die and a black die.

Sample space S =
{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) }



X = sum of both dice

function with X(1,1) = 2, X(1,2) = 3, …, X(6,6)=12

From Random Variables to Events

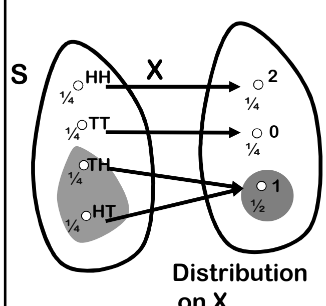
Note that each event in the induced distribution corresponds to some event in the original one.

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

Pr(A) = Pr(X=a) = Pr({t ∈ S | X(t)=a})

Two Coins Tossed

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



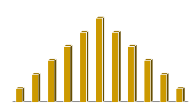
Pr(X = a) =
Pr({t ∈ S| X(t) = a})

Pr(X = 1)
= Pr({t ∈ S| X(t) = 1})
= Pr({TH, HT}) = ½

Two dice

I throw a white die and a black die.

Sample space S =
{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) }



X = sum of both dice

Pr(X = 7) = $\frac{1}{6} = \frac{6}{36}$

Definition: Expectation

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

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

X is a function on the sample space S

X has a prob. distribution on its values

(assuming X takes values in the naturals)

Quick check: $\sum_{t \in S} Pr(t)\ X(t) = \sum_{k} k\ Pr[X = k]$

A Quick Calculation…

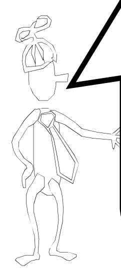
What if I flip a coin 3 times? What is the expected number of heads?

E[X] = (1/8)×0 + (3/8)×1 + (3/8)×2 + (1/8)×3 = 1.5

But Pr[X = 1.5] = 0

Moral: don't always expect the expected.
Pr[X = E[X]] may be 0 !

Type Checking



P(B) B must be an event

E(X) X must be a R.V.

cannot do P(R.V.) or E(event)

Operations on R.V.s

You can sum them, take differences,
or do most other math operations…

E.g., (X + Y)(t) = X(t) + Y(t)

(X*Y)(t) = X(t) * Y(t)

$(X^Y)(t) = X(t)^{Y(t)}$

Random variables and expectations allow us to give elegant solutions to problems that seem really really messy…

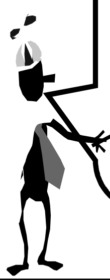
If I randomly put 100 letters into 100 addressed envelopes, on average how many letters will end up in their correct envelopes?

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

Pretty messy with direct counting…

The new tool is called
"Linearity of Expectation"

Linearity of Expectation



If Z = X+Y, then

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

Even if X and Y are not independent

$$E[Z] = \sum_{t \in S} Pr[t]\, Z(t)$$

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

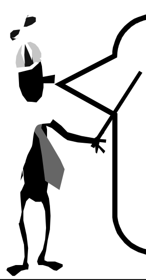
$$= \sum_{t \in S} Pr[t]\, X(t) + \sum_{t \in S} Pr[t]\, Y(t))$$

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

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

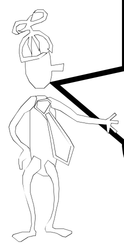
Expectation of a Sum of r.v.s
= Sum of their Expectations

even when r.v.s are not independent!

Expectation of a Product of r.v.s
false in general = Product of their Expectations

only when r.v.s are independent!

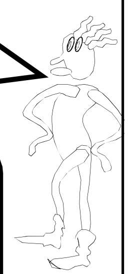
Independence for r.v.s



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?

Let's test our Linearity of Expectation chops…



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

Hmm…

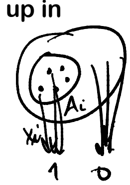
$\sum_k$ k Pr(exactly k letters end up in correct envelopes)

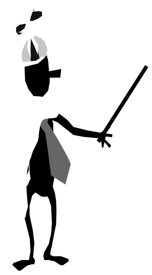
= $\sum_k$ k (…aargh!!…)

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





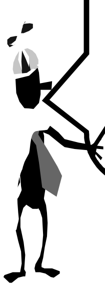
$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

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

We are asking for E[Z] $= \sum_i E[X_i]$

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

So E[Z] = 1



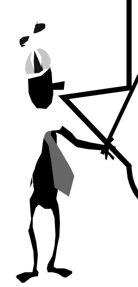
So, in expectation, 1 letter will be in the same correct envelope

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



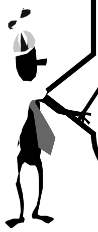
General approach:

View thing you care about as expected value of some RV

Write this RV as sum of simpler RVs

Solve for their expectations and add them up!

Example #2



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

We could do this by summing

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

$= n.p$

But now we know a better way!

Linearity of Expectation!

Let X = number of heads when n independent coins of bias p are flipped

$X = \sum x_i$

Break X into n simpler RVs:

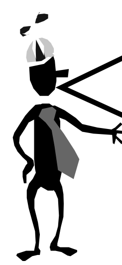
$$X_i = \begin{cases} 1 \text{ if the } i^{th} \text{ coin is } \cancel{\text{tails}} \text{ heads} \\ 0 \text{ if the } i^{th} \text{ coin is } \cancel{\text{heads}} \text{ tails} \end{cases}$$

$E[X] = E[\Sigma_i X_i] = \Sigma_i E[X_i] = \Sigma_i p = np$

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

Σ_k k Pr(exactly k collisions)

$= \Sigma_k$ k (…aargh!!!!…)

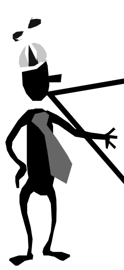


Use linearity of expectation

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

X = number of pairs of people with the same birthday

E[X] = ?



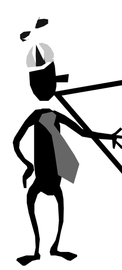
X = number of pairs of people with the same birthday

E[X] = ?

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

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

$E[X_{jk}] = (1/366)\ 1 + (1 - 1/366)\ 0$
$= 1/366$



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

$E[X_{jk}] = (1/366)\ 1 + (1 - 1/366)\ 0$
$= 1/366$

$E[X] = E[\Sigma_{1 \le j < k \le m} X_{jk}]$
$= \Sigma_{1 \le j < k \le m} E[X_{jk}]$
$= m(m-1)/2 \times 1/366$

E.g., setting m = 23, we get
E[# pairs with same birthday]
= 0.691…

How does this compare to
Pr[at least one pair has same birthday] ?

You pick a number n ∈ [1..6].
You roll 3 dice.
If any of them match n, you win $1.
Else you pay me $1. Want to play?

X_i = 1 if i-th die matches, 0 otherwise

Expected number of dice that match:
$E[X_1+X_2+X_3] = 1/6+1/6+1/6 = ½$.

BUT not same as Pr(at least one die matches).

Pr(at least one die matches) = 1 – Pr(none match)
= 1 – $(5/6)^3$ = 0.416.

Back to Lin. of Exp.

Using linearity of expectations in unexpected places…

A Puzzle

10% of the surface of a sphere is colored green, and the rest is colored blue. Show that now matter how the colors are arranged, it is possible to inscribe a cube in the sphere so that all of its vertices are blue.

Solution

Pick a random cube. (Note: any particular vertex is uniformly distributed over surface of sphere).

Let X_i = 1 if i^{th} vertex is blue, 0 otherwise

Let $X = X_1 + X_2 + \ldots + X_8$

$E[X_i] = P(X_i=1) = 9/10$
$E[X] = 8*9/10 > 7$
So, must have some cubes where X = 8.

The general principle we used here

Show the expected value of some random variable is “high”

Hence, there must be an outcome in the sample space where the random variable takes on a “high” value.

(Not everyone can be below the average.)

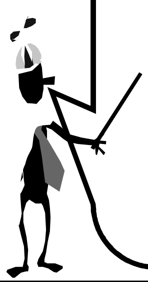
called “the probabilistic method”

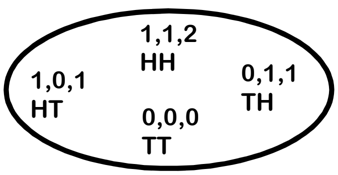
Some supplementary material

linearity of expectation:
it holds even when the
random variables are
not independent

Linearity of Expectation

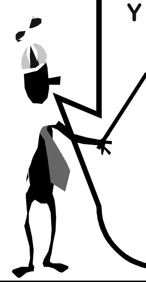
E.g., 2 fair flips:
X = heads in 1st coin,
Y = heads in 2nd coin
Z = X+Y = total # heads
What is E[X]? E[Y]? E[Z]?

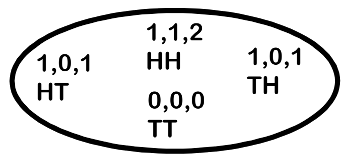




Linearity of Expectation

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]?





from random variables
to events
and vice versa

From Random Variables to Events

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

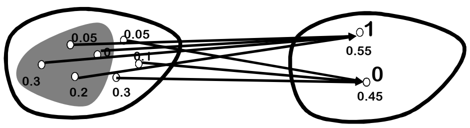
$$\Pr(A) = \Pr(X=a) = \Pr(\{t \in S \mid X(t)=a\})$$

From Events to Random Variables

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

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

$\boxed{E[X_A] = 1 \times Pr(X_A = 1) = Pr(A)}$



How about multiplying expectations?

Multiplication of Expectations

A coin is tossed twice.

X_i = 1 if the i^{th} toss is heads and 0 otherwise.

$E[\,X_i\,] = 1/2$

$E[\,X_1 X_2\,] = E[\,X_1\,]\,E[\,X_2\,] = 1/4$

$E[XY] = E[X]E[Y]$ if they are independent

Multiplication of Expectations

Consider a single toss of a coin.

X = 1 if heads turns up and 0 otherwise.

Set $Y = 1 - X$

$E[X] = E[Y] = 1/2$

X and Y are not independent

$E[X\,Y] \neq E[X]\,E[Y]$

since $X\,Y = 0$