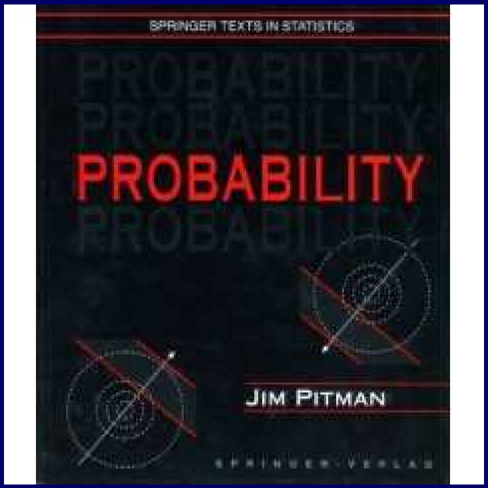
Introduction to probability

Stat 134 FAll 2005

Berkeley



Lectures prepared by:

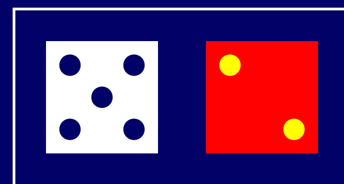
Elchanan Mossel
Yelena Shvets

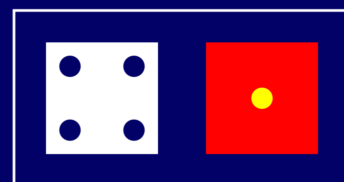
Follows Jim Pitman's book:

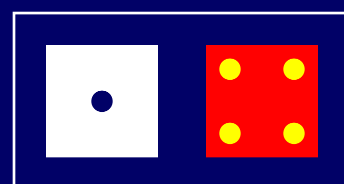
Probability

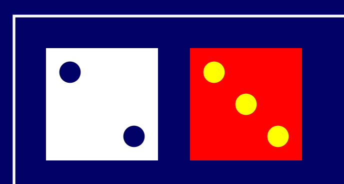
Section 3.1

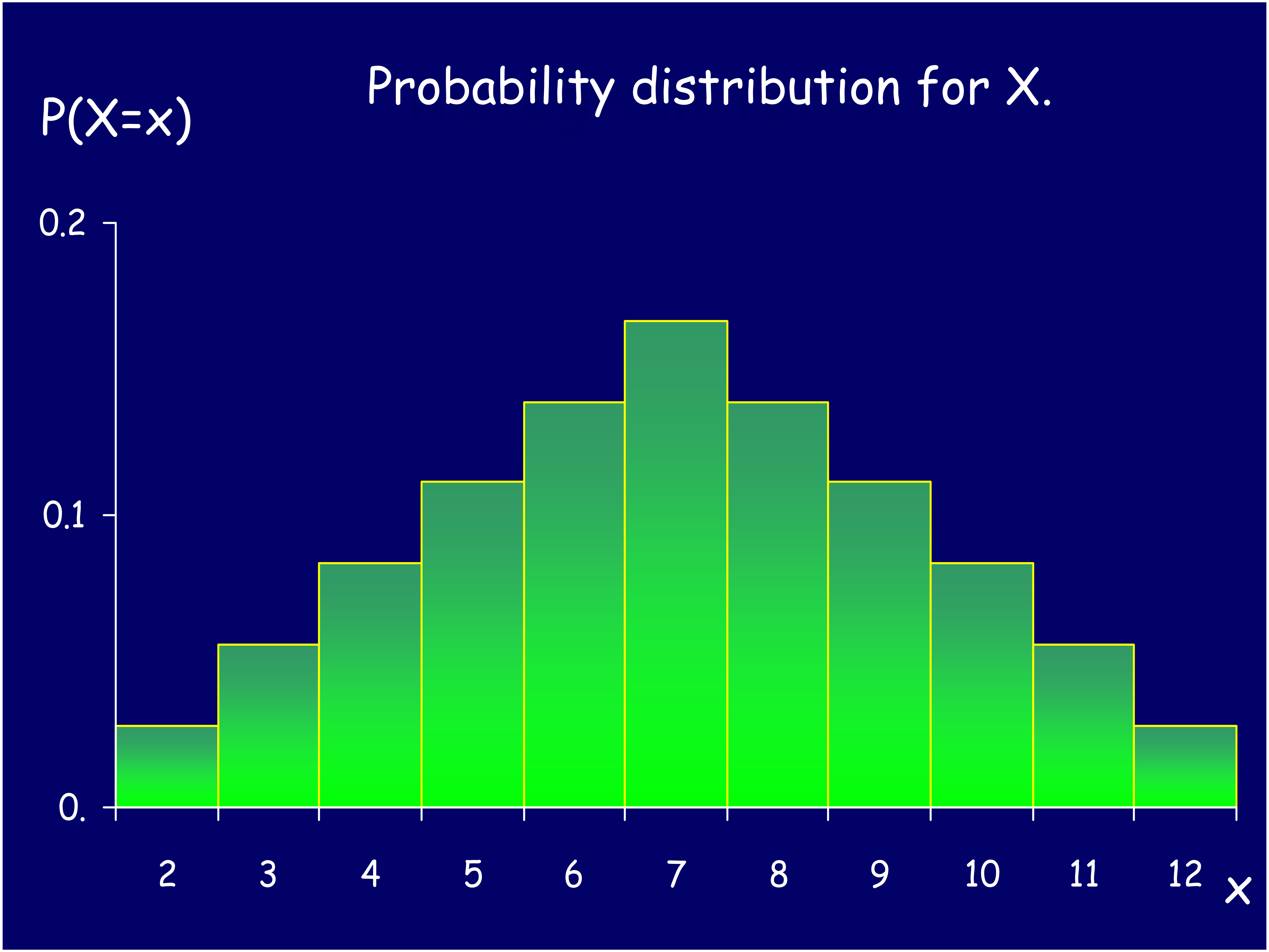
X is the sum of two dice.

$X($  $) = 7$

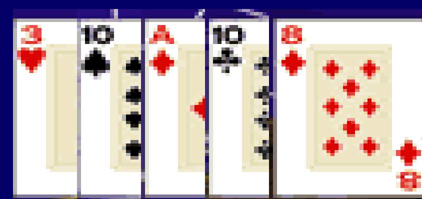
$X($  $) = 5$

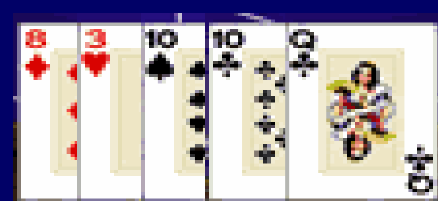
$X($  $) = 5$

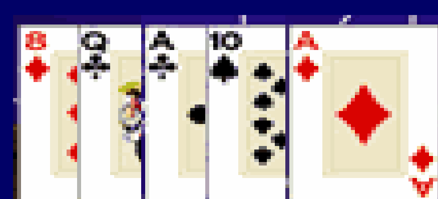
$X($  $) = 5$

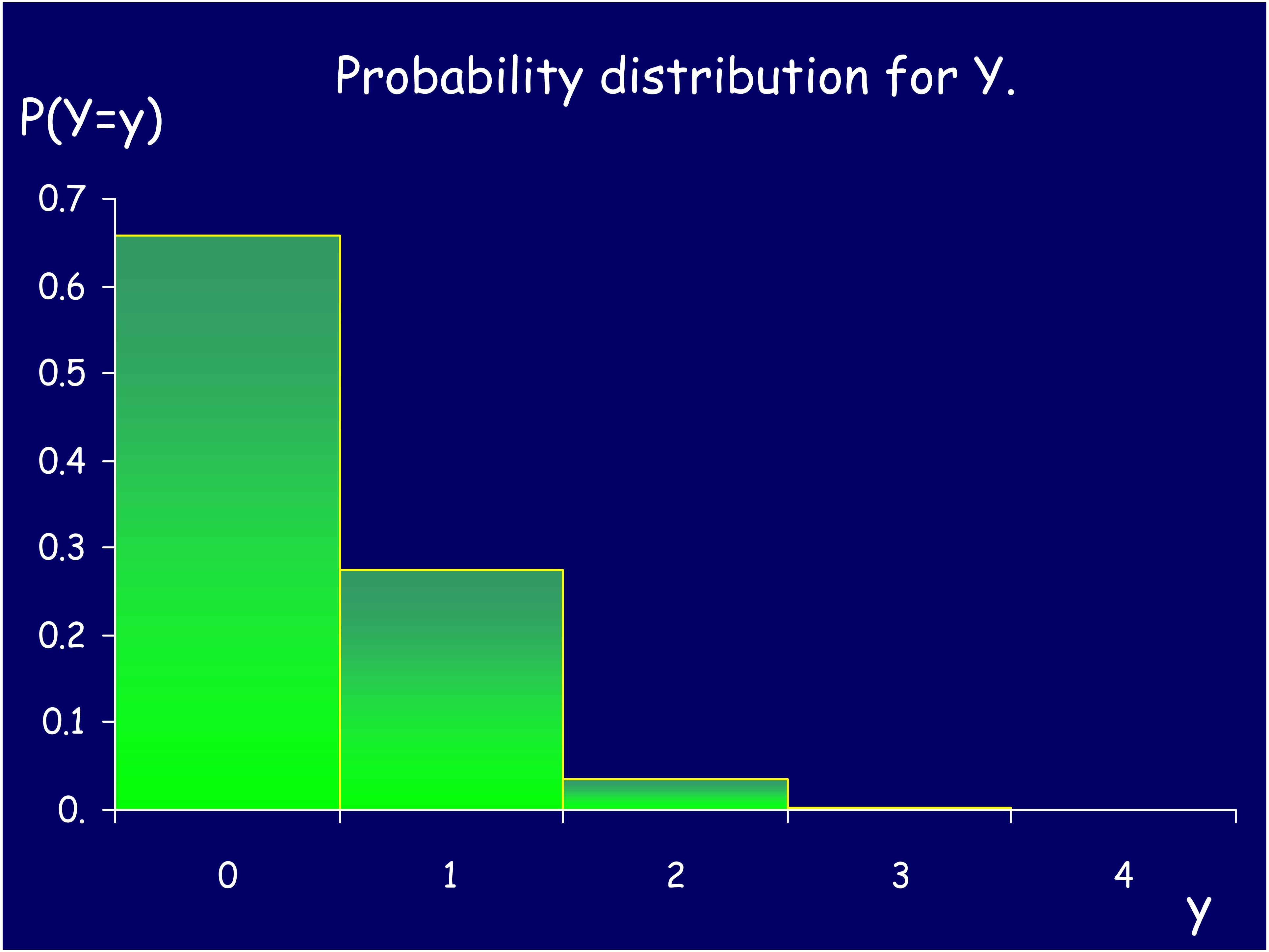

Probability distribution for X.
P(X=x)
0.2
0.1
0.
2
3
4
5
6
7
8
9
10
11
12
x

Y is the number of aces in a poker hand.

Y () =1

Y () =0

Y () =2



Probability distribution for Y.
P(Y=y)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
0
1
2
3
4
y

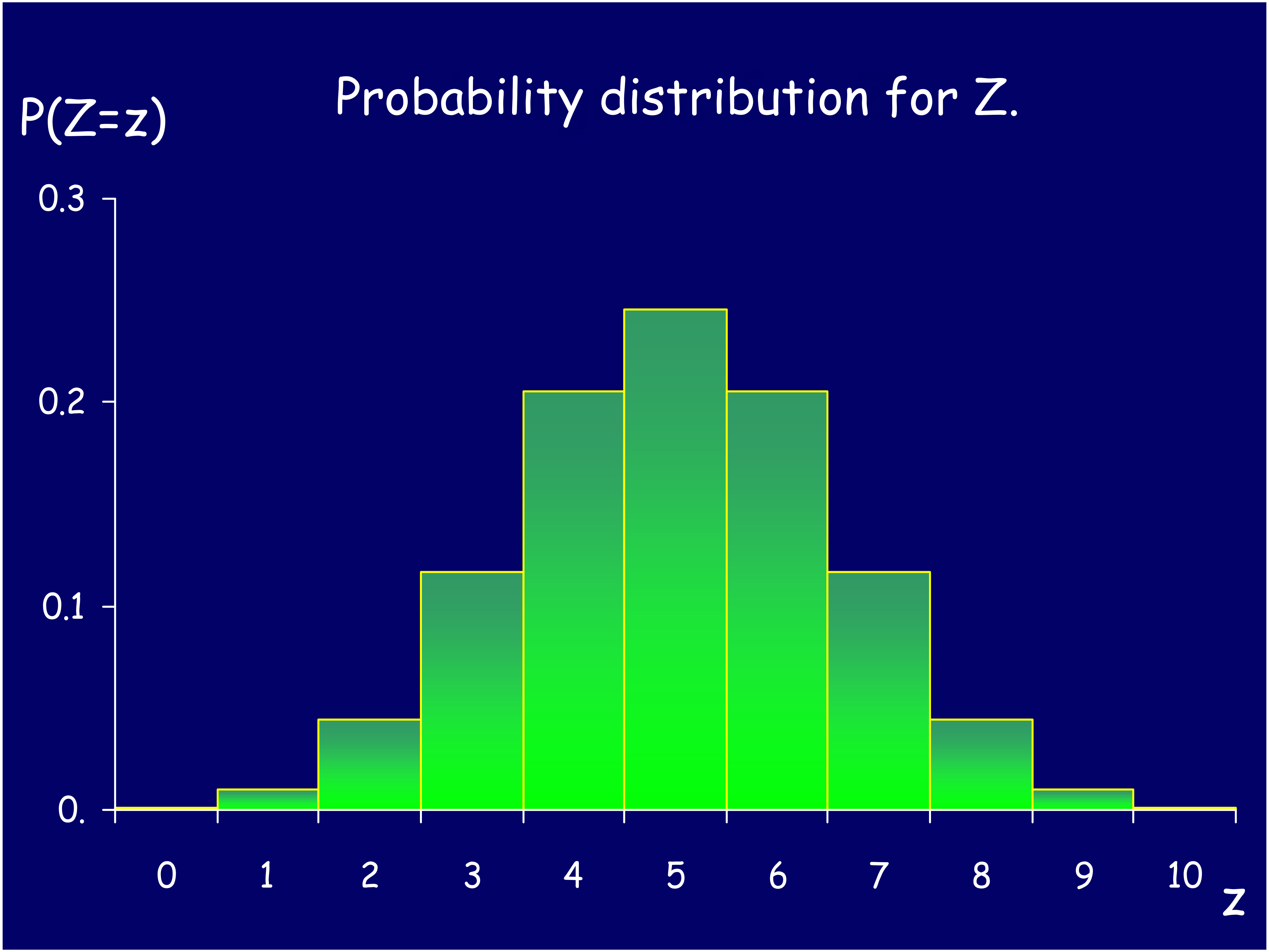
Z is the number of heads in 10 coin tosses.

Z()=5

Z()=5

Z()=7

Z()=6



Probability distribution for Z.
P(Z=z)
0.3
0.2
0.1
0.
0
1
2
3
4
5
6
7
8
9
10
z

Random Variables

A Random Variable is a symbol representing a quantity that is determined by an outcome in a probability space.

It represents a numerical value.

Algebraic Variable - X=3 (which means X is 3)
Random Variable - P(X=3) (which means X has a probability of being 3).

Random Variables

Consider Z, the number of heads in 10 coin tosses.

The range of Z is the set of all possible values that Z can take:

$$\{0,1,2,3,4,5,6,7,8,9,10\}.$$

Random Variables

What is the domain of Z?

In other words what determines the value that Z takes?

This is Ω = the space of outcomes. In this case, an outcome is a particular sequence of coin tosses.

Z() =5

Ω has 2^{10} = 1024 elements.

Random Variables

In other words, a random variable is like a function that takes in the real world and spits out a number in its range.

Ex. X = the # on a dice roll (Range: 1-6)

$W = X^2$ (1,4,9,16,25,36)

Y = the # of ppl who are awake (0-~70)

Z = the # of ppl born today (non-neg. Int)

V = the kg. of grain harvested this year in Canada (Real no.)

There is some probability that the random variable will equal every value in its range.

Random Variables

Note that for every random variable X,

$$\sum_x P(X=x) = 1,$$

where x ranges over all possible values that X can take.

All the possible values that X can take and the probability that X is any one of those values is called the distribution of X.

Several Random Variables

Given two RV X and Y defined on the same outcome space Ω, we can consider a random pair (X,Y).

$$(X,Y)(\omega)=(X(\omega),Y(\omega)).$$

Then event (X,Y)=(x,y) can be written as:

$$\{\omega:(X,Y)(\omega)=(x,y)\}$$

$$= \{\omega:X(\omega)=x\} \cap \{\omega:Y(\omega)=y\}$$

This event has probability $P(X = x \text{ and } Y = y)$ where x and y are in the ranges of X and Y.

Example:

Let X be the sum of two dice and Y be the minimum.

X,Y	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2,1	3,1	4,1	5,1	6,1	7,1
⚁	3,1	4,2	5,2	6,2	7,2	8,2
⚂	4,1	5,2	6,3	7,3	8,3	9,3
⚃	5,1	6,2	7,3	8,4	9,4	10,4
⚄	6,1	7,2	8,3	9,4	10,5	11,5
⚅	7,1	8,2	9,3	10,4	11,5	12,6

Joint Distributions

The distribution of (X,Y) is called the joint distribution of X and Y:

$$P(x,y) = P(X=x, Y=y),$$

satisfying

$P(x,y) \geq 0$ and $\sum_{\text{all }(x,y)} P(x,y) = 1.$

We can build a probability distribution table for the X and Y in the previous example:

P(X,Y)	1	2	3	4	5	6
2	1/36	0	0	0	0	0
3	2/36	0	0	0	0	0
4	2/36	1/36	0	0	0	0
5	2/36	2/36	0	0	0	0
6	2/36	2/36	1/36	0	0	0
7	2/36	2/36	2/36	0	0	0
8	0	2/36	2/36	1/36	0	0
9	0	0	2/36	2/36	0	0
10	0	0	0	2/36	1/36	0
11	0	0	0	0	2/36	0
12	0	0	0	0	0	1/36

Joint Distributions

Question:

Suppose you know the distribution of X and the distribution of Y separately, does this determine their joint distribution?

Several Random Variables

Example:

X is the result of the first draw from the box containing 1,2,3,4,5,6;

&

Y – the result of the second draw after the first ticket has been replaced.

Joint Distribution

X \ Y	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

P(X,Y)

Several Random Variables

Example:

X is the result of the first draw from the box containing 1,2,3,4,5,6;

&

Z – the result of the second draw without replacing the first ticket.

Joint Distribution

X \ Z	1	2	3	4	5	6
1	0	1/30	1/30	1/30	1/30	1/30
2	1/30	0	1/30	1/30	1/30	1/30
3	1/30	1/30	0	1/30	1/30	1/30
4	1/30	1/30	1/30	0	1/30	1/30
5	1/30	1/30	1/30	1/30	0	1/30
6	1/30	1/30	1/30	1/30	1/30	0

P(X,Z)

Joint Distribution

X \ Z	1	2	3	4	5	6	dist. X (row sums)
1	0	1/30	1/30	1/30	1/30	1/30	P(X=1)=1/6
2	1/30	0	1/30	1/30	1/30	1/30	P(X=2)=1/6
3	1/30	1/30	0	1/30	1/30	1/30	P(X=3)=1/6
4	1/30	1/30	1/30	0	1/30	1/30	P(X=4)=1/6
5	1/30	1/30	1/30	1/30	0	1/30	P(X=5)=1/6
6	1/30	1/30	1/30	1/30	1/30	0	P(X=6)=1/6
dist. Z (column sums)	P(Z=1)=1/6	P(Z=2)=1/6	P(Z=3)=1/6	P(Z=4)=1/6	P(Z=5)=1/6	P(Z=6)=1/6	

Joint Distributions

Question:

Suppose you know the distribution of X and the distribution of Y separately, does this determine their joint distribution?

Answer:

It does not …

Marginal Probabilities

$$P(X=x) = \sum_{y} P(X=x, Y=y)$$

$$P(Y=y) = \sum_{x} P(X=x, Y=y)$$

Joint Distribution

x \ y	1	2	3	4	5	6	dist. X (row sums)
1	1/36	1/36	1/36	1/36	1/36	1/36	P(X=1)=1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	P(X=2)=1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	P(X=3)=1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	P(X=4)=1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	P(X=5)=1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	P(X=6)=1/6
dist. Y (column sums)	P(Z=1)=1/6	P(Z=2)=1/6	P(Z=3)=1/6	P(Z=4)=1/6	P(Z=5)=1/6	P(Z=6)=1/6	

Random Variables with the Same Distribution

Random Variables X and Y have the same or identical distribution if they have the same range and for every value x in their range

$$P(X=x) = P(Y=x).$$

Equality of Random Variables

Having the same distribution is not enough. If we consider the function analogy, two random variables are equal if they spit out the same number for every status of the real world.

Of course if they are equal, they will have the same distribution.

Recall:

X is the result of the first draw from the box $\boxed{1,2,3,4,5,6}$;

Y – the result of the second draw after the first ticket has been replaced;

Z – the result of the second draw without replacing the first ticket

Draws from a box.

Observe that X, Y & Z all have the same distribution: uniform on {1,2,3,4,5,6}.

Yet they are not equal as variables:

If X = 3 then Z cannot equal 3.

(Y and Z are not even defined on the same space of outcomes…)

Change of Variables Principle

If X and Y have the same distribution then so do $g(X)$ and $g(Y)$, for any function g. For example:

X^2 has the same distribution as Y^2.

Equality of Random Variables

Random variables X and Y defined over the same outcome space are equal, written X=Y, iff

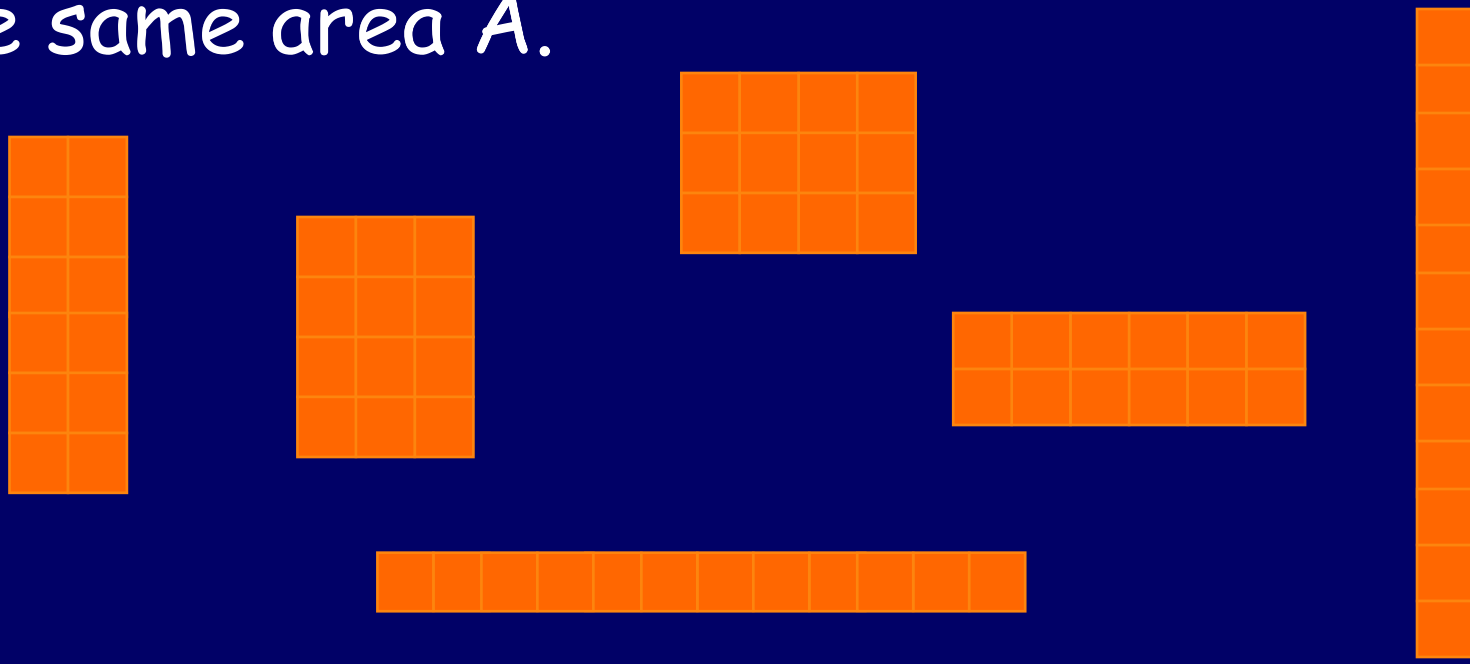
X= Y for every outcome in Ω.

In particular, P(X=Y) = 1.

Equal RV's

Example:

Suppose we have a box with tiles each having the same area A.



Let X be the height of the tiles, Y be their width and Z = A/X. Then Z=Y as random variables.

Probabilities of Events Determined by X and Y

The probability that X and Y satisfy some condition is the sum of P(x,y) over all pairs (x,y) which satisfy that condition.

$$P(X < Y) = \sum_{(x,y):x<y} P(x,y) = \sum_{\text{all } x} \sum_{y:y>x} P(x,y)$$

$$P(X = Y) = \sum_{(x,y):x=y} P(x,y) = \sum_{\text{all } x} P(x,x)$$

Functions of Random Variables

We can add two random variables:

If X is the no. on a 6 sided die

Y is {3 if Professor Mossel is wearing a blue shirt today, 0 if not}

Then X+Y is then sum of the two individual random variable values.

How many ways can X+Y = 0? 1? 4? 8?

Distribution of a function of X and Y.

Distribution of any function of X and Y can be determined from the joint distribution:

$$P(f(X,Y)=z) = \sum_{(x,y):f(x,y)=z} P(x,y)$$

Functions of Random Variables

Let X,Y represent the results of two draws with replacement from a box with (1, 2,…,6).

Let Z=max(X,Y). We say Z=f(X,Y).

	1	2	3	4	5	6	P(Z=z)
1	1	2	3	4	5	6	**P(Z=1)=1/36**
2	2	2	3	4	5	6	P(Z=2)=3/36
3	3	3	3	4	5	6	P(Z=3)=5/36
4	4	4	4	4	5	6	P(Z=4)=7/36
5	5	5	5	5	5	6	P(Z=5)=9/36
6	6	6	6	6	6	6	P(Z=6)=11/36

Conditional Distribution given an event.

For any even A, and any random variable Y, the collection of conditional probabilities

$$P(Y \in B \mid A) = P(Y \in B, A)/P(A)$$

defines the conditional distribution of Y given A.

Conditional Distribution Given A.

We can compute $P(Y \in B | X \in A)$ by partitioning B:

$$P(Y \in B \mid X \in A) = P(Y \in B, X \in A) / \sum_y P(Y=y, X \in A).$$

Hence, if Y has a finite range this distribution is specified by $P(Y=y|A)$ for $y \in$ range of Y.

We can let $A : X = x$, for any random variable X defined over the same space of outcomes.

Conditional Distribution of Y Given X=x.

For every value of x in the range of X, as y varies over the range of Y the probabilities

$$P(Y=y|X=x)$$

define the conditional distribution of Y given X=x.

This is related to conditional probabilities of events, as if we know more information (X=x), some events are more/less likely to happen just as random variables are more/less likely to take on certain values.

Rules of Conditional Probability

- Non-negative: $P(Y \in B|A) \geq 0$ for $A,B \subseteq \Omega$.

- Additive: if $B = B_1 \sqcup B_2 \sqcup \ldots \sqcup B_n$

$$P(Y \in B|A) = P(Y \in B_1 |A) + P(Y \in B_2 |A) + \ldots + P(Y \in B_n |A).$$

- Sums to 1: $P(Y \in \Omega |A) = 1$

Multiplication Rule

If the marginal distribution of X and the conditional distribution of Y given X=x are known, then we can find the joint distribution of X and Y:

$$P(X=x, Y=y) = P(X=x)\,P(Y=y|X=x).$$

Independence

Suppose that for any x and y

$$P(Y = y \mid X = x) = P(Y=y).$$

This means that knowing what value X takes does not influence the probabilities for Y so values of X and Y do not depend on each other. Note that this must work for all possible x and y in the ranges of X and Y.

This property defines independence.

Independence

Now suppose we know the marginal distribution of X and the conditional distribution of Y given X=x.

Then the joint density is:

$$P(X=x, Y=y) = P(X=x)\, P(Y=y|X=x).$$

If X and Y are independent then

$$P(X=x, Y=y) = P(X=x)\, P(Y=y).$$

This is an equivalent definition of independence.

Independence

Suppose we know that for X and Y

$$P(X=x, Y=y) = P(X=x)\,P(Y=y).$$

Then
$$\begin{aligned} P(Y=y \mid X=x) &= P(X=x, Y=y)\,/P(X=x) \\ &= (P(X=x)\,P(Y=y))/P(X=x) \\ &= P(Y=y). \end{aligned}$$

This property does not always hold!

$$P(X=x, Y=y) \neq P(X=x)\, P(Y=y)$$

for general X and Y.

Dependence

Example: draws without replacement from a box of {1,2,3,4,5,6}.

X \ Z	1	2	3	4	5	6
1	0	1/30	1/30	1/30	1/30	1/30
2	1/30	0	1/30	1/30	1/30	1/30
3	1/30	1/30	0	1/30	1/30	1/30
4	1/30	1/30	1/30	0	1/30	1/30
5	1/30	1/30	1/30	1/30	0	1/30
6	1/30	1/30	1/30	1/30	1/30	0

$$P(X=1, Y=1) = 0$$

$$\neq$$

$$P(X=1)\,P(Y=1) = 1/36$$

Sequence of 10 coin tosses

Suppose we denote a head by 1 and a tail by 0.

Then a sequence of 10 coin tosses can be represented as a sequence of zeros and ones:

$$\omega = 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 1.$$

Ω has $2^{10} = 1024$ elements

0000000000	1000000000
0000000001	1000000001
0000000010	1000000010
0000000011	1000000011
0000000100	1000000100
0000000101	1000000101
0000000110	1000000110
:	:
0111111111	1111111111

Sequence of 10 coin tosses

We can define 10 new RV's over Ω in the following way:

outcome: 0100011011;

$X^1 = 0, X^2 = 1, X^3 = 0, \ldots, X^{10} = 1.$

X^i represents the outcome of the i^{th} coin toss.

Sequence of 10 coin tosses

We can show that X^i's are pair-wise independent Bernoulli(½) variables by using the first definition of independence.

$$P(X^i = 1 \mid X^j = 1) = P(X^i = 1 \mid X^j = 0) = P(X^i = 1) = \tfrac{1}{2};$$

In fact it's enough to look at ω^1 and ω^{10}, by symmetry.

$P(X^{10}=1 \mid X^{1}=0) = 1/2 = P(X^{10}=1)$

0000000000	1000000000
0000000001	1000000001
0000000010	1000000010
0000000011	1000000011
0000000100	1000000100
0000000101	1000000101
0000000110	1000000110
:	:
0111111111	1111111111

Ω has $2^{10} = 1024$ elements

Joint Distribution of Several Random Variables.

The distribution of $(X_1, X_2, \ldots, X_n)$ is called the **joint distribution** of $X_1, X_2, \ldots, X_n$:

$$P(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n),$$

satisfying

$$P(x_1, x_2, \ldots, x_n) \geq 0 \quad \text{and}$$

$$\sum_{\text{all } (x_1, x_2, \ldots, x_n)} P(x_1, x_2, \ldots, x_n) = 1.$$

Independence of Several Random Variables

The second definition of independence is easily generalized to the case of several RV's.

We say $X_1, X_2, \ldots, X_n$ are independent if

$$P(X_1=x_1,\ldots,X_n=x_n) = P(X_1=x_1)\times\ldots\times P(X_n=x_n).$$

Functions of Independent RV's are Independent

If $X_1, X_2, \ldots, X_n$ are independent then so are the random variables

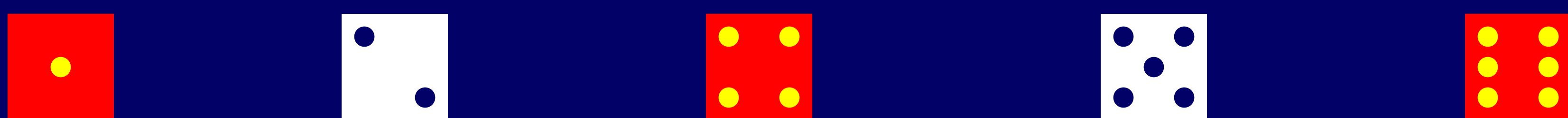
$$Y_1 = f_1(X_1)\ , \ldots\ , Y_n = f_n(X_n).$$

For any functions $f_1, \ldots, f_n$ defined on the range of $X_1, \ldots, X_n$.

Example:

X_1, X_2, X_3,X_4,X_5 are the numbers on 5 dice, and f(x) = 0 if x is even and f(x) = 1 if x is odd. If we let $Y_i = f(X_i)$.,Then Y_1,Y_2,Y_3,Y_4,Y_5 are independent.

X_1 X_2 X_3 X_4 X_5

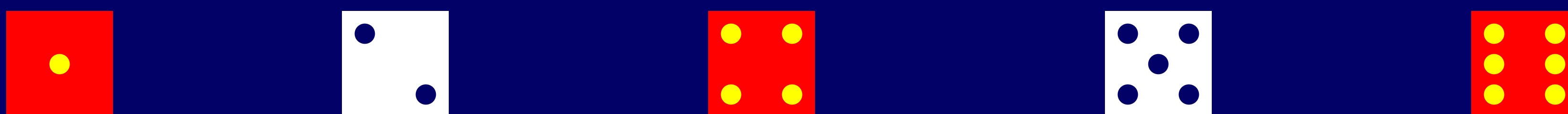


$Y_1 = 1$ $Y_2 = 0$ $Y_3 = 0$ $Y_4 = 1$ $Y_4 = 0$

Disjoint Blocks of Independent RV's are Independent

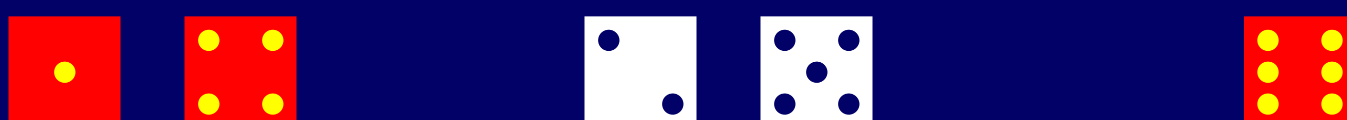
Let's consider the 5 dice again.

X_1 X_2 X_3 X_4 X_5



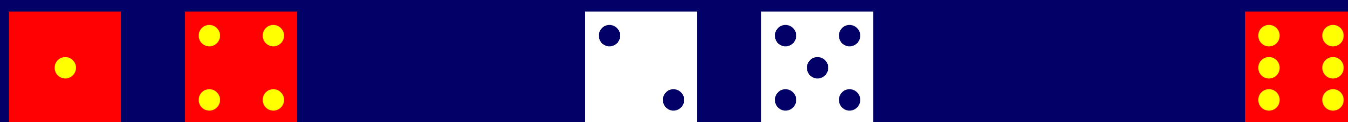
Suppose we group them into random vectors.

(X_1, X_3) (X_2, X_4) (X_5)



Disjoint Blocks of Independent RV's are Independent

(X_1, X_3) (X_2, X_4) (X_5)



These new random vectors are independent.

Disjoint Blocks of Independent RV's are Independent

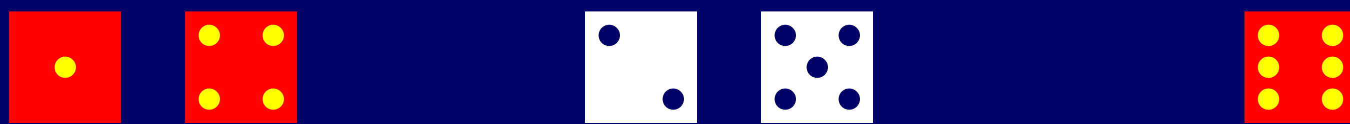
If $X_1,\ldots,X_n$ are independent.

And $B_1,\ldots,B_k$ are disjoint subsets of positive integers such that
$B_1 \cup B_2 \ldots \cup B_k = \{1,2,\ldots,n\}$.

Then the random vectors
$(X_i : i \in B_1), (X_i : i \in B_2), \ldots, (X_i : i \in B_k)$
are independent.

Disjoint Blocks of Independent RV's are Independent

(X_1, X_3) (X_2, X_4) (X_5)



In this example

$$B_1 = \{1,3\};\ B_2 = \{2,4\};\ B_3 = 5.$$

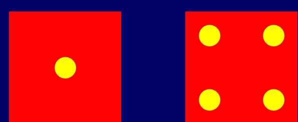
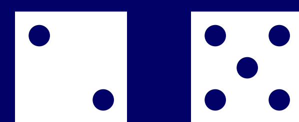
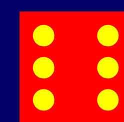
Functions of Disjoint Blocks of Independent RV's are Independent

If $X_1, X_2, \ldots, X_5$ are independent then so are the random variables

$$Y_1 = X_1 + X_3^2; \quad Y_2 = X_2 X_4; \quad Y_3 = X_5$$

Disjoint Blocks of Independent RV's are Independent

(X_1, X_3) (X_2, X_4) (X_5)

$Y_1 = X_1 + X_3^2$
$= 17$

$Y_1 = X_2 X_4$
$= 10$

$Y_3 = X_5$
$= 6$