PRIME INTUIT
Finishing School

Random Variables

Recap: List of Topics

Descriptive Statistics	✓ Different types of data ✓ Different types of plots ✓ Measure of centrality and Spread
Probability Theory	✓ Counting, Sample Specs, events, axioms ✓ Conditional Probability (3 Laws) ✓ RVs Expectation and Distributions ✓ Sampling strategies
Inferential Statistics	✓ Interval Estimators ✓ Hypothesis testing (z-test, t-test) ✓ ANOVA, Chi-square test ✓ Linear Regression

Introduction to Random Variables

Objectives or Topics

What are ***random variables*** *?*

What is a ***probability mass function*** *?*

What are some ***standard probability mass functions*** *?*

What are ***Expectations and Variance*** *(and some of their properties) ?*

What ***is probability density function*** *?*

What is a ***normal distribution*** *?*

Introduction to Random Variables

Recap

Experiments, sample spaces, events

Axioms and Laws of Probability

Now we will focus on numerical quantities associated with the outcomes of experiments

Introduction to Random Variables

Mapping outcomes to R

In board games we are more concerned about the sum and not the outcomes which lead to the sum

$\Omega \to \{\}$

$11 \subset I \subset R$

Int Real no.

Question: What is the probability that the sum will be 10

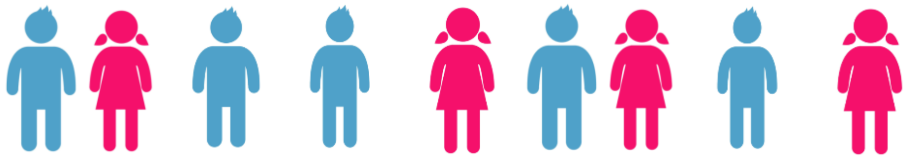
$\Omega \to R$

	Sum
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

Introduction to Random Variables

Mapping outcomes to R

Ω



What is the probability that a students CGPA is 4.5 ?

Questions:

What is the probability that the size of the farm is less than 2 acres?

What is the probability that that total yield is greater than 1 ton ?

R

$CGPA$

3 3.5 4 4.5 5 5.5

Height, Weight, Vit B3, shoes size, etc

Introduction to Random Variables

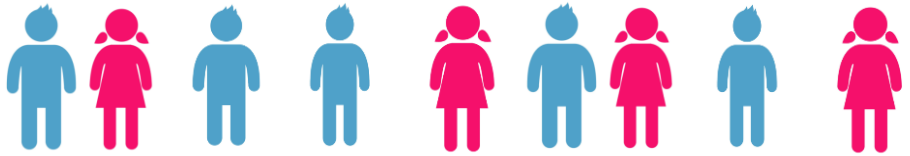
Mapping outcomes to R

Questions:

What is the probability that an employee has 2 children?

What is the probability that an employee's salary is greater than 50K?

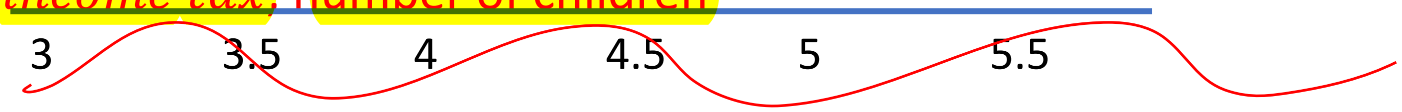
Ω



Ω: All employees of the organisation?

R

Number of years of experience, number of projects, salary, income tax, number of children



Introduction to Random Variables

Mapping outcomes to R

Randomly select a farm

Ω

Ω : All farms in the state

Questions:

What is the probability that the size of the farm is less than 2 acres?

What is the probability that that total yield is greater than 1 ton ?

R: Size of the farm, total yield, soil moisture, water content

Introduction to Random Variables

Mapping outcomes to R

X : Function — Ω Domain $\rightarrow$ R Range

(Random Variable) — (Could be a subset of R)

Ω	R
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

A random variable is a function from a set of possible outcomes to the set of real numbers

Introduction to Random Variables

Mapping outcomes to R

Ω → R

Domain   Range

X1 : height → R

X2 : weight → R

X3 : CGPA → R

Ω
Students
Domain

Multiple functions (random variables) are possible for the given domain (sample space)

Ω	R
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

Introduction to Random Variables

Notation

X(n1, n2): Sum of the numbers on 2 dice (n1 + n2)

Y(student): height of the student

X Y

Unlike functions we don't write brackets and arguments!

Ω	R
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

Introduction to Random Variables

Questions

X

What are the values that the random variable can take ?

(discrete or continuous)

What are the probabilities of the values that the random variable can take ?

(we will return back to this later)

Ω	R
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

PRIME INTUIT
Finishing School

Types of Random Variables

Discrete (finite or countably infinite)

The sum of the numbers on two dice

The outcome of a single die

The number of tosses after which a heads appears (countably infinite)

The number of children that an employee has

The number of cars in an image

Ω		R
(1, 1)	2	
(1, 2) (2, 1)	3	
(1, 3) (2, 2) (3, 1)	4	
(1, 4) (2, 3) (3, 2) (4, 1)	5	
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6	
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7	
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8	
(3, 6) (4, 5) (5, 4) (6, 3)	9	
(4, 6) (5, 5) (6, 4)	10	
(5, 6) (6, 5)	11	
(6, 6)	12	

Types of Random Variables

Continuous

The amount of rainfall in Mysore

The temperature of a surface

The density of a liquid

The height of a student

The haemoglobin level of a patient

Ω	R
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

Probability Mass Function

Probability of Mass function

What are the probabilities of the values that the random variables can take ?

$R \rightarrow P(R)$

$2 \rightarrow 2/6$

$3 \rightarrow 3/6$

4

6

Probability Mass Function

Assigning probabilities

$X: \Omega \rightarrow R$

{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

Question: What is the probability that the value o fthe random variable will be X

P (X = x) ?

$\forall x \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

P(X = x) → [0, 1]

An assignment of probabilities to all possible values that a discrete RV can take is called the distribution of the discrete random variable

For discrete RVs we can think of the distribution as a table

X	P(X = x)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Probability Mass Function

PRIME INTUIT Finishing School

Assigning probabilities

$X: \quad \Omega \quad \rightarrow \quad R$

{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

P(X = x) → [0, 1]

Tips:
Think of the event corresponding to X = x

Once we know this event (subset of sample space) we know how to compute P[X = x]

$Px(x)$ = P[X = x] = P $(w \in \Omega : X(w) = x)$

Probability distribution of the random variable X ,
(probability distribution or probability mass function)

(1, 1)

(1, 2) (2, 1)

(1, 3) (2, 2) (3, 1)

(1, 4) (2, 3) (3, 2) (4, 1)

(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)

(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)

(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)

(3, 6) (4, 5) (5, 4) (6, 3)

(4, 6) (5, 5) (6, 4)

(5, 6) (6, 5)

(6, 6)

For discrete RVs we can think of the distribution as a table

X	P(X = x)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

PRIME INTUIT
Finishing School

Probability Mass Function

Probability Mass Function

$Px(x)$

Probability Mass Function (PMF)

Probability Distribution

Distribution

(1, 1)
(1, 2) (2, 1)
(1, 3) (2, 2) (3, 1)
(1, 4) (2, 3) (3, 2) (4, 1)
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)
(3, 6) (4, 5) (5, 4) (6, 3)
(4, 6) (5, 5) (6, 4)
(5, 6) (6, 5)
(6, 6)

Properties of PMF

Probability Mass Function

$\mathrm{P}x(x) \geq 0$

$\mathrm{P}x(x) = P(X = x) = P(\{w \in \Omega : X(w) = x\}) \geq 0$?

$\Sigma_{x \in R_x} \mathrm{P}x(x) = 1$

$R_x \subset R$

(the set of values that the RV can take)

(the support of the RV)

PRIME INTUIT
Finishing School

Properties of PMF

Proof:

$$\Sigma_{x \in R_x} \mathrm{P}x(x) = 1$$

$$R_x \subset R$$

$$\Sigma_{x \in R_x} \mathrm{P}x(x) = \Sigma_{x \in R_x} P(X = x)$$

RHS is the sum of the probabilities of disjoint events which partition Ω

Ω	R
(1, 1)	2
(1, 2) (2, 1)	3
(1, 3) (2, 2) (3, 1)	4
(1, 4) (2, 3) (3, 2) (4, 1)	5
(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	6
(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	7
(2, 6) (3, 5) (4, 4) (5, 3) (6, 2)	8
(3, 6) (4, 5) (5, 4) (6, 3)	9
(4, 6) (5, 5) (6, 4)	10
(5, 6) (6, 5)	11
(6, 6)	12

Discrete distributions

Probability Mass Functions for discrete random variables

Recap

X :	Ω	$\rightarrow$	R
Function	**Domain**		**Range**
(Random Variable)			(Could be a subset of R)

Distribution of a random variable

An assignment of probabilities to all possible values that a discrete RV can take

(can be tedious even in simple cases)