UNIVERSITY OF LINCOLN

SCHOOL OF MATHEMATICS AND PHYSICS

# MTH1005
# PROBABILITY AND STATISTICS

*Semester B*
*Lecture 4*
*(20/2/2024)*

***Danilo Roccatano***

**Office:** INB 3323
**Email:** droccatano@lincoln.ac.uk

# SUMMARY OF CONDITIONAL PROBABILITIES

- **Definition of conditional probabilities for two events (E,F) in S**

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

- **Multiplicative rule for a set of events ($E_1$… $E_n$) in S**

$$P(E_1 E_2 E_3 \ldots E_n) = P(E_1) \frac{P(E_1 \cap E_2)}{P(E_1)} \frac{P(E_1 \cap E_2 \cap E_3)}{P(E_1 \cap P(E_2)} \cdots \frac{P(E_1 \cap E_2 \cdots \cap E_n)}{P(E_1 \cap E_2 \cdots \cap E_{n-1})}$$

# SUMMARY OF CONDITIONAL PROBABILITIES

**- Law of total probability for a set of events ($E_1$… $E_n$) in S**

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) P(E_i)$$

**- Bayes' Formula for a set of events $E_i$**

$$P(E_i \mid A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(A \mid E_i) P(E_i)}{\sum_{j=1}^{n} P(A \mid E_j) P(E_j)}$$

# INDEPENDENT EVENTS

**Definition:** Two experiments are independent if the result of one cannot in any way affect the possible results of the other.

**Definition:** Two events (E, F ) are independent if the probability that one of them occurs is in no way in influenced by whether or not the other has occurred.

$$P(E) = P(E \mid F) = P(E \mid \bar{F}),$$
$$P(F) = P(F \mid E) = P(F \mid \bar{E}).$$

put in a different way this means that

$$P(E \cap F) = P(E)P(F)$$

*the probability of E and F occurring is just the product of the probability of E occurring and the probability of F occurring.*

# INDEPENDENT EVENTS

**Theorem:** Two events (E, F ) are independent if and only if

$$P(E \cap F) = P(E)P(F)$$

For more than two events things become a bit more restrictive.

**Theorem:** The events $E_1, E_2, E_3, \cdots E_n$ are said to be mutually independent if for every subset

$E'_1, E'_2, E'_3, \ldots E'_r, r \leq n$

$$P(E'_1 \cap E'_2 \cap E'_3 \cap \cdots \cap E'_r) = P(E'_1) \cdot P(E'_2) \cdot P(E'_3) \cdots P(E'_r)$$

# INDEPENDENT EVENTS

The idea of independent processes will be extremely important as we move forward.

**We will use the idea that repetitions of experiments constitute independent processes very often.**

**Note that this is more or less an assumption of the frequentist definition of probability.**

# INDEPENDENT EVENTS: EXAMPLE

Consider the compound experiment of throwing two fair coins.
The sample space is

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

Define two events

*A = 'the first coin is a head' = {(H, H), (H, T )}*
*B = 'the second coin is a tail' = {(H, T ), (T , T )}*

P (A)= |A|/|S| = 2/4 = ½ and
P (B)= |B|/|S| = 2/4 = ½

Now $P(A \cap B) = P(H, T) = |A \cap B|/|S| = 1/4 = P(A) * P(B)$.

so A and B are independent.

# INDEPENDENT EVENTS: EXAMPLE

But consider C = 'both coins are heads' = {(H, H)}

**$P(C) = |C|/|S| = 1/4$**

now

**$P(B \cap C) = P(\emptyset) = |B \cap C|/|S| = 0/4 \neq P(B) * P(C) = (1/2)(1/4)$,**
**$P(A \cap C) = P(\emptyset) = |A \cap C|/|S| = 0/4 \neq P(A) * P(C) = (1/2)(1/4)$,**

so B and C are not independent. A is also not independent of C .

In general, it is possible for all pairs of events to be independent, but the complete set of events not to be.

# TABULAR PRESENTATION OF CONDITIONAL PROBABILITIES

It can sometimes be handy to view conditional probabilities using a tabular representation of relative frequencies - this can also be how real data arrives to us.

# TABULAR PRESENTATION OF CONDITIONAL PROBABILITIES

## THE BUS EXAMPLE

A town has three bus routes, A, B, and C. During rush hour there are twice as many buses on the A route as on B or C.

Over time, it has been observed that at a crossroads, where the routes converge, the buses run more than 5 minutes late 1/2, 1/5, 1/10 of the time.

*If an inspector at the crossroads finds that the first bus, he sees is more than five minutes late, what it the chance that it is a route B bus?*

# TABULAR PRESENTATION OF CONDITIONAL PROBABILITIES

If we go back to the bus problem we did the last lecture, but instead we consider what we'd expect if **1000** buses in total ran through the town, we'd end up with something like

| | A | B | C | total |
|---|---|---|---|---|
| Late | 250 | 50 | 25 | 325 |
| Not late | 250 | 200 | 225 | 675 |
| Total | 500 | 250 | 250 | 1000 |

## TABULAR PRESENTATION OF CONDITIONAL PROBABILITIES

**Let L** = 'a bus is late'

| | A | B | C | total |
|---|---|---|---|---|
| Late | 250 | 50 | 25 | 325 |
| Not late | 250 | 200 | 225 | 675 |
| Total | 500 | 250 | 250 | 1000 |

The conditional probabilities can easily be read off, for instance

$$\mathbf{P\ (B \mid L) = P\ (B \cap L)/P\ (L) = 50/325 = 2/13}$$

This is of course exactly equivalent to our pen and paper solution.

# TABULAR PRESENTATION OF CONDITIONAL PROBABILITIES

| | A | B | C | total |
|---|---|---|---|---|
| Late | 250 | 50 | 25 | 325 |
| Not late | 250 | 200 | 225 | 675 |
| Total | 500 | 250 | 250 | 1000 |

**Marginal Probabilities**

This is because of the law of total probability.

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) P(E_i)$$

## TABULAR PRESENTATION OF CONDITIONAL PROBABILITIES

| | A | B | C | total |
|---|---|---|---|---|
| Late | 250 | 50 | 25 | 325 |
| Not late | 250 | 200 | 225 | 675 |
| Total | 500 | 250 | 250 | 1000 |

Another example

$$P(L) = P(L \cap A) + P(L \cap B) + P(L \cap C) = 250 + 50 + 25 = 325$$

or **vertically** we have

$$P(A) = P(A \cap L) + P(A \cap \overline{L})$$

note that that last relationship is a useful and general one.
It is a special case of the law of total probability.

# TABULAR BAYES' THEOREM: yet another TABULAR EXAMPLE

A doctor is trying to decide if a patient has one of three diseases **$d_1$, $d_2$, or $d_3$.**

- **Two tests** are to be carried out, each of which results in a positive (+) or a negative (−) outcome.
- There are four possible test patterns **++, +−, −+,** and **−−**.
- National records have indicated that, for 10,000 people having one of these three diseases, the distribution of diseases and test results are as in the table in the next slide.

# TABULAR BAYES' THEOREM: yet another TABULAR EXAMPLE

| Disease | number | + + | + – | – + | – – |
|---|---|---|---|---|---|
| d1 | 3215 | 2110 | 301 | 704 | 100 |
| d2 | 2125 | 396 | 132 | 1187 | 410 |
| d3 | 4660 | 510 | 3568 | 73 | 509 |
| Total | 10000 | | | | |

We can use this data to estimate **$P(d_1)$, $P(d_2)$, $P(d_3)$ -** these are called prior probabilities, and the conditional probabilities like

$$P\ (+- \mid d_1) = 301/3215 = 0.094$$

# TABULAR BAYES' THEOREM: yet another TABULAR EXAMPLE

What the doctor wants though is

*the probability a patient has disease $d_i$ given the results of the tests.*

These are the ***Bayes' or inverse, or posterior probabilities.***

We can compute them using Bayes' formula and we'll get results like

- Bayes' Formula

$$P(E_i \mid A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(A \mid E_i)P(E_i)}{\sum_{j=1}^{n} P(A \mid E_j)P(E_j)}$$

| | $d_1$ | $d_2$ | $d_3$ |
|---|---|---|---|
| $++$ | .700 | .131 | .169 |
| $+-$ | .075 | .033 | .892 |
| $-+$ | .358 | .604 | .038 |
| $--$ | .098 | .403 | .499 |

**Ex.**

**$P(d_1 \mid ++) = P(++ \mid d_1)P(d_1)/P(++) = 0.700$ .**

# TABULAR BAYES' THEOREM: A TABULAR EXAMPLE

The values in the table are **P($d_i$, ++)** etc. Judicious use of these posterior probabilities can be used to inform decision:

| | $d_1$ | $d_2$ | $d_3$ |
|---|---|---|---|
| $++$ | .700 | .131 | .169 |
| $+-$ | .075 | .033 | .892 |
| $-+$ | .358 | .604 | .038 |
| $--$ | .098 | .403 | .499 |

If the test result came back ++ then the posterior probability P($d_1$, ++) = 0.7 and we'd suspect that $d_1$ was the culprit.

# UPDATING INFORMATION SEQUENTIALLY

- In the last example, we updated the likelihood of a patient having a particular disease based on extra information in the form of the test results.
- The example should also give a hint about how we can update our ideas about a system as new information comes in.
- this is why the original probabilities are called priors, and the reversed conditional probabilities are called posterior probabilities.
- *We'll use this type of method in the computing lab later as a simple form of machine learning.*

# RANDOM VARIABLES

# RANDOM VARIABLES

In the next  sessions we will define, discuss and begin to use random variables.

Understand the elements of a random variable.
Be able to define and use

- Discrete vs continuous variables
- Range and domain of the variables
- Probability (Mass/Density) Function
- Cumulative Distribution Function Improper integrals

# RANDOM VARIABLES

Random variables are a tool that let us work with more complicate situations using still using a lot of tools that we know from algebra or calculus.

Unlike a normal variable, it doesn't make sense to say 'x = some value'.

*For example:*

- How many miles to a gallon does my car do?
- How quickly do I run a mile ?
- How long will it take for a website to load?
- What will be the average daily temperature in two days?

# RANDOM VARIABLES

Random variables inherently can take on a variety of values, and only discussing a statistic distribution of their values makes sense. Every time we tried something (took a measurement) we would get a slightly different answer.

Most idealisations of real-life experiments would correspond to random variables!

We'll see

- what random variables are
- how they can link abstract probability distributions to real-life experiments

## STRING LOGIC

The random variable is a series of 'mappings', from events to probabilities to numerical values.

We

- split the sample space into events,
- we assign each event a probability,
- we assign each event a unique numerical value.

the random variable is the whole of that package.

# STRING LOGIC

Lets consider rolling a normal fair six sided die.

We want to quantify the number of spots on the topmost face of die after a single roll.

If we had an infinite amount of information we could determine which face will come up before the roll was made. But as we don't we can only work out the probabilities of the different faces coming up.

With those different faces coming up, and their probabilities we associate numbers.

## STRING LOGIC

These steps can be done in any order, but often it makes sense to assign the values we want to the random variable.

We'll call our random variable Z.

We'll mix up some mathematical notation and common sense

$$Z = \begin{cases} 1 & \text{if 'the die lands with 1 spot on the top face'} \\ 2 & \text{if 'the die lands with 2 spots on the top face'} \\ 3 & \text{if 'the die lands with 3 spots on the top face'} \\ 4 & \text{if 'the die lands with 4 spots on the top face'} \\ 5 & \text{if 'the die lands with 5 spots on the top face'} \\ 6 & \text{if 'the die lands with 6 spots on the top face'} \end{cases}$$

the statements if 'the die lands with 1 spot on the top face' hopefully sound a lot like events to you.

# STRING LOGIC

S = {1, 2, 3, 4, 5, 6}
where {1} is the outcome that the die lands with 1 spot on the top face, {2} is the outcome that the die lands with 2 spots up etc.

So looking again at Z

$$Z = \begin{cases} 1 & \text{if 'the die lands with 1 spot on the top face'} \\ 2 & \text{if 'the die lands with 2 spots on the top face'} \\ 3 & \text{if 'the die lands with 3 spots on the top face'} \\ 4 & \text{if 'the die lands with 4 spots on the top face'} \\ 5 & \text{if 'the die lands with 5 spots on the top face'} \\ 6 & \text{if 'the die lands with 6 spots on the top face'} \end{cases}$$

we could replace the text 'the die lands with 1 spot on the top face' with {1}, the event that the die lands with 1 spot upwards, and get

# STRING LOGIC

$$Z = \begin{cases} 1 & \text{if } \{1\} \\ 2 & \text{if } \{2\} \\ 3 & \text{if } \{3\} \\ 4 & \text{if } \{4\} \\ 5 & \text{if } \{5\} \\ 6 & \text{if } \{6\} \end{cases}$$

## STRING LOGIC

Now what is the probability that we get those different values of Z ?

We get Z = 1 if the event {1}, or 'the die lands with 1 spot on the top face' occurs.

You should know how to associate a probability to an event occurring.

In this case where all the outcomes {1, 2, 3, 4, 5, 6} can be argued to be equally likely, the probability of any one of them occurring must be 1/|S|, one over the number of elements in the sample space, in this case 1/6 .

So the probability that Z= 1 , which is that {1} occurred is 1/6

# STRING LOGIC

we'll de ne another function, the probability (mass) function of Z .

$$p_Z(z) = P(Z = z) = \begin{cases} 1/6 & \text{if } \{1\} \\ 1/6 & \text{if } \{2\} \\ 1/6 & \text{if } \{3\} \\ 1/6 & \text{if } \{4\} \\ 1/6 & \text{if } \{5\} \\ 1/6 & \text{if } \{6\} \end{cases}$$

and our random variable ***Z*** is fully defined.

# EXAMPLE OF 6-sided DIE

we have our sample space

S = {1, 2, 3, 4, 5, 6}

the values of our random variable, and their associated probabilities

$$Z = \begin{cases} 1 & \text{if } \{1\} \\ 2 & \text{if } \{2\} \\ 3 & \text{if } \{3\} \\ 4 & \text{if } \{4\} \\ 5 & \text{if } \{5\} \\ 6 & \text{if } \{6\} \end{cases}, p_Z(z) = P(Z = z) = \begin{cases} 1/6 & \text{if } \{1\} \\ 1/6 & \text{if } \{2\} \\ 1/6 & \text{if } \{3\} \\ 1/6 & \text{if } \{4\} \\ 1/6 & \text{if } \{5\} \\ 1/6 & \text{if } \{6\} \end{cases}$$

we must have that each possible value of the random variable is unambiguously associated with an event and a probability.
It should be clear that this is the case above.

# DEFINITION OF DISCRETE RANDOM VARIABLE

**Definition** Let $X$ be a discrete random variable that can take on only the values $x_1, x_2, x_3, \ldots, x_n$ with respective probabilities $p_X(x_1), p_X(x_2), p_X(x_3), \ldots, p(x_n)$. Then, if

$$\sum_{x \in X} p_X(x) = 1,$$

$X$ is a **discrete random variable.**
The function $p(x) = P\{X = x\}$ is called the probability (mass) function of the variable $X$.

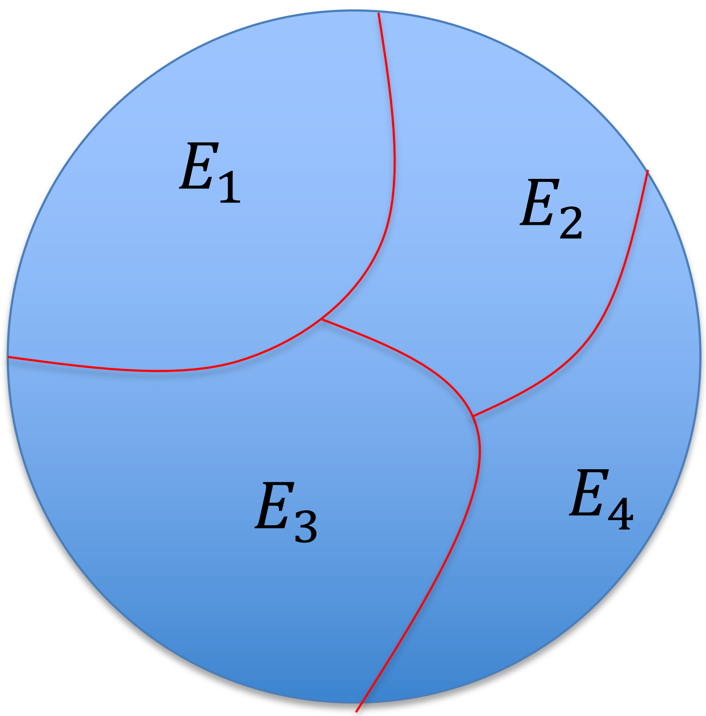
**NOTE**: Other notations are sometimes used for the probability (mass) function, $m(i) or\ p_i$.

It is also conventional that random variables are given a capital letter, and normally are chosen from the end of the alphabet).

# DISCRETE RANDOM VARIABLES

**Definition:** Consider an experiment, with outcome set S , split into *n* mutually *exclusive* and *exhaustive* events $E_1, E_2, E_3, \ldots, E_n$.
A variable, X say, which can assume exactly n numerical values each of which corresponds to one and only one of the given events is called a random variable.

Schematically the mutually exclusive and exhaustive events look like




***Event types:***

- Exclusive: no overlap.
- Exhaustive: all of S is covered by them.

## DISCRETE RANDOM VARIABLES

Here our outcome set is split into 4 mutually exclusive events and exhaustive. So to associate a random variable (call it **X** ) with this sample space we could have something like

```
X = 1, corresponding to event E₁
```

$$X = 1, \text{ corresponding to event } E_1$$

$$X = 2, \text{ corresponding to event } E_2$$

$$X = 3, \text{ corresponding to event } E_3$$

$$X = 4, \text{ corresponding to event } E_4$$

and we know how to calculate the probability of events.

## STRING LOGIC

The random variable is a series of 'mappings', from events to probabilities to numerical values.

We

- split the sample space into events,
- we assign each event a probability,
- we assign each event a unique numerical value.

the random variable is the whole of that package.

## *Check list*

- is the sample space well defined?
- are the events mutually exclusive?
- do the events cover all the sample space?
- are the probabilities of the events defined
- are values of the variable clearly assigned one-to-one to the possible events?

# EXAMPLE

*Define a valid random variable to describe the number of girls in families with 2 children, assuming that the likelihoods of boys or girls is equal and independent.*

First, we de ne our sample space

*S = {(B, B), (B, G), (G, B), (G, G)}*

We define our variable G using the number of girls in the family:

$$G = \begin{cases} 0 & \text{if } \{(B, B)\} \\ 1 & \text{if } \{(B, G), (G, B)\} \\ 2 & \text{if } \{(G, G)\} \end{cases}$$

## EXAMPLE

and associate our probability (mass) function

$$p_G(x) = \begin{cases} p_G(0) = P(G=0) = P(\{(B,B)\}) = & \frac{1}{4} \\ p_G(1) = P(G=1) = P(\{(B,G),(G,B)\}) = & \frac{1}{2} \\ p_G(2) = P(G=2) = P(\{(G,G)\}) = & \frac{1}{4} \end{cases}$$

We check that

$$\sum_{g \in G} p_G(g) = 1.$$

Here *g* indicates the possible values of G, so

$$\sum_{g \in G} p_G(g) = p_G(0) + p_G(1) + p_G(2) = 1/4 + 1/2 + 1/4 = 1$$

## EXAMPLE 2

Construct a random variable that counts the total number of spots upwards when two fair 6-sided dice are thrown. Our sample space can be the 36 elements of

$$S = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\}$$

where *i* is the number of spots on the first die, and j the second die.

## EXAMPLE 2

How we need to think of a set of mutually exclusive and exhaustive events that would correspond to the sum of the spots. We could write this in set builder notation

$$E_k = \{(i, j) : i + j = k;\ i, j = 1, 2, 3, 4, 5, 6\}$$

and we can write these out in full

$$
\begin{aligned}
E_2 &= \{(1,1)\} \\
E_3 &= \{(1,2),(2,1)\} \\
E_4 &= \{(1,3),(2,2),(3,1)\} \\
E_5 &= \{(1,4),(2,3),(3,2),(4,1)\} \\
E_6 &= \{(1,5),(2,4),(3,3),(4,2),(5,1)\} \\
E_7 &= \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \\
E_8 &= \{(2,6),(3,5),(4,4),(5,3),(6,2)\} \\
E_9 &= \{(3,6),(4,5),(5,4),(6,3)\} \\
E_{10} &= \{(4,6),(5,5),(6,4)\} \\
E_{11} &= \{(5,6),(6,5)\} \\
E_{12} &= \{(6,6)\}
\end{aligned}
$$

## EXAMPLE 2

and we assign the value i + j to the random variable X if the event $E_k$ occurs.

$$X = \begin{cases} 2 & \text{if } E_2 \\ 3 & \text{if } E_3 \\ \vdots & \\ 11 & \text{if } E_{11} \\ 12 & \text{if } E_{12} \end{cases}$$

## EXAMPLE 2

and the probability distribution associated with X is similarly defined

$$p_X(x) = P(X = x) = P(E_x) = \begin{cases} 1/36 & \text{if } X = 2 \\ 2/36 & \text{if } X = 3 \\ 3/36 & \text{if } X = 4 \\ 4/36 & \text{if } X = 5 \\ 5/36 & \text{if } X = 6 \\ 6/36 & \text{if } X = 7 \\ 5/36 & \text{if } X = 8 \\ 4/36 & \text{if } X = 9 \\ 3/36 & \text{if } X = 10 \\ 2/36 & \text{if } X = 11 \\ 1/36 & \text{if } X = 12 \end{cases}$$

The lower case x gives particular values of X, and that collection is the range of X - the values that X can take on:

$$R_X = 2, 3, 4, \ldots, 10, 11, 12$$

in that language the domain of the variable X is the sample set S.

# STRING LOGIC: SUMMARY

*Check list for random variable*

- is the sample space well defined?
- are the events mutually exclusive?
- do the events cover all the sample space?
- are the probabilities of the events defined
- are values of the variable clearly assigned one-to-one to the possible events?

# (CUMULATIVE) DISTRIBUTION FUNCTION

The cumulative distribution function of a discrete random variable, X , is

$$F_X(a) = P\{X \leq a\} = \sum_{x \leq a} p_X(x)$$

it is an alternative way of describing the probabilities of different events.

Sometimes it is more convenient to use $F_X$ than the probability function $p_X$ .

It provided a convenient way of describing continuous random variables.

## EXAMPLE

Let us imagine we have a 4 sided die.

What is the cumulative distribution function for the random variable
*X = 'what result appears when a 4 sided die is thrown once'*

First we build our random variable.
All the probabilities can be assumed equally likely, and we will assign a value of the number of spots to our variable X.

We have a sample space S={1,2,3,4}, and

$$X = \begin{cases} 1 & \text{if } \{1\} \\ 2 & \text{if } \{2\} \\ 3 & \text{if } \{3\} \\ 4 & \text{if } \{4\} \end{cases}$$

# EXAMPLE

We define our probability function

$$p_X(1) = P(X = 1) = P(\{1\}) = \frac{1}{4}$$
$$p_X(2) = P(X = 2) = P(\{2\}) = \frac{1}{4}$$
$$p_X(3) = P(X = 3) = P(\{3\}) = \frac{1}{4}$$
$$p_X(4) = P(X = 4) = P(\{4\}) = \frac{1}{4}$$

or more sensibly

$$p_{X(x)} = \frac{1}{4} : x = 1,2,3,4$$

We see that

$$\sum_{x \in X} p_{X(x)} = 1$$

The events that define X are exhaustive and mutually exclusive.
There is a one-to-one map between the values of the probability (mass) function. So X is a random variable, with $p_X(x)$ its probability mass function.

## EXAMPLE

Using our definition of the cumulative distribution function

$$F(a) = P\{X \leq a\} = \sum_{x \leq a} p(x)$$

we have

$$F_X(a) = \begin{cases} 0 & a < 1 \\ 1/4 = & 1 \leq a < 2 \\ 2/4 = & 2 \leq a < 3 \\ 3/4 = & 3 \leq a < 4 \\ 1 = & a >= 4 \end{cases}$$

# Definition

$$F_X(a) = P\{X \leq a\} = \sum_{x \leq a} p_X(x) \text{ for } -\infty < a < \infty$$

if this is true, then because $p_X(x)$obeys the axioms of probability, the following will hold:

$F_X(a)$ must be 0 as $a \rightarrow -\infty$ and 1 as $a \rightarrow \infty$

$F_X(a)$ must be monotonically increasing

$F_X(a) \quad$ *must be right continuous*

## CUMULATIVE DISTRIBUTION FUNCTION FOR CONTINUOUS RANDOM VARIABLES

Cumulative distributions provide a good way of describing continuous random variables.

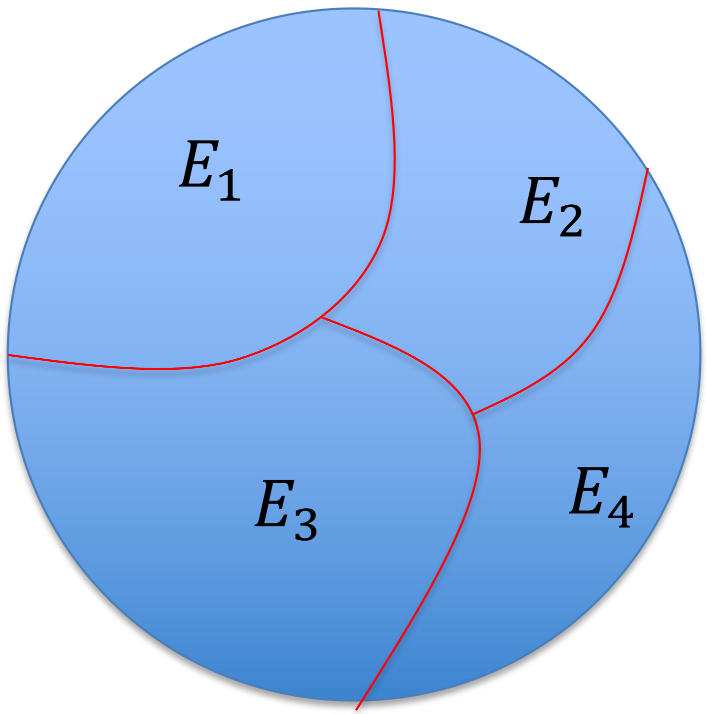
The cumulative distribution function of a discrete random variable, X, is

$$F_X(a) = P\{X \leq a\} = \sum_{x \leq a} p_X(x)$$

## DISCRETE RANDOM VARIABLES

**Definition:** Consider an experiment, with outcome set S , split into *n* mutually *exclusive* and *exhaustive* events $E_1, E_2, E_3, \ldots, E_n$.
A variable, X say, which can assume exactly n numerical values each of which corresponds to one and only one of the given events is called a random variable.

Schematically the mutually exclusive and exhaustive events look like




***Event types:***

- Exclusive: no overlap.
- Exhaustive: all of S is covered by them.

# (CUMULATIVE) DISTRIBUTION FUNCTION

The cumulative distribution function of a discrete random variable, X , is

$$F_X(a) = P\{X \leq a\} = \sum_{x \leq a} p_X(x)$$

it is an alternative way of describing the probabilities of different events.

Sometimes it is more convenient to use $F_X$ than the probability function $p_X$ .

It provided a convenient way of describing continuous random variables.

# CONTINUOUS RANDOM VARIABLES AND CONDITIONAL PROBABILITY

## INFINITE, AND CONTINOUS PROBABILITY SPACES

Everything up to this point deals with probability spaces with a finite sample space.

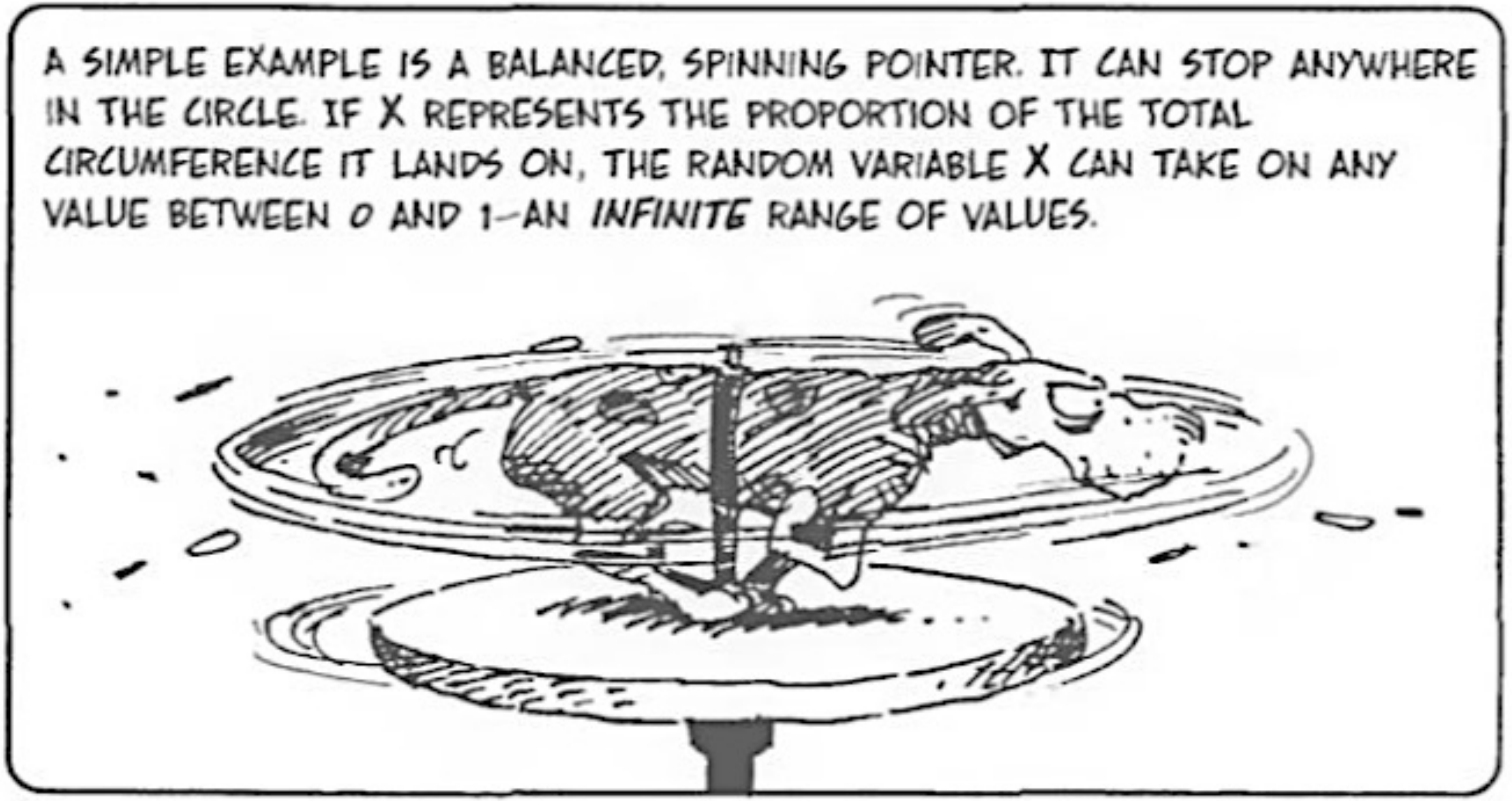
Here we mention two other cases

- when the sample space has discrete points, but an infinite number of them (all integers, for instance),

- or a prevailing situation when the sample space can take on a continuous range of values.

# CONTINOUS PROBABILITY SPACES

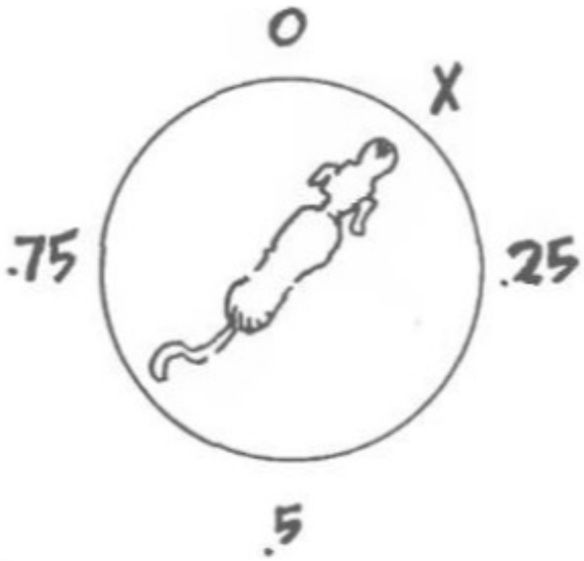
Consider a spinner - schematically a circle of unit circumference and a pointer




*From: Gonick & Smith The cartoon guide to Statistics*

# CONTINOUS PROBABILITY SPACES

If we give the spinner a whirl, the pointer will be pointing somewhere a distance x along the circumference.




It seems reasonable that every value $0 \leq x < 1$ *of* the distance between the pointer and the mark on the spinner is equally likely to occur. That means that the sample space is the interval S = [0, 1).

We would like to have a probability model for which every value of the sample space is equally likely. We'll call the result of a spin X for now, and later we'll see that this is a *continuous random variable*.

# CONTINOUS PROBABILITY SPACES

In a similar way to before we must have

$$P(0 \leq X < 1) = 1$$

It is also the case that we expect the probability of a reading in the top half of the spinner is equal in likelihood to one in the lower half,

$$P\left(0 \leq X < \frac{1}{2}\right) = P\left(\frac{1}{2} \leq X < 1\right) = \frac{1}{2}$$

More generally, if we consider an event, E= [a, b], we'd like

$$P(a \leq X < b) = b - a$$

for every a and b .

# CONTINOUS PROBABILITY SPACES

We can satisfy

$$P(a \leq X < b) = b - a$$

for every a and b for the event E= [a, b] by a formula of the form

$$P(E) = \int_E f(x)dx$$

and *f(x) is* constant function with value 1.

We call f(x) *the density function* of X.

This is the generalisation of the discrete case we saw earlier:

$$P(E) = \sum_{i \epsilon E} P(i)$$

# CONDITIONAL CONTINUOUS PROBABILITIES

If we look at a process that has a density function *f(x)*, and if *E* is an event. We define a conditional density function by

$$f(x \mid E) = \begin{cases} f(x)/P(E) & \text{if } x \in E, \\ 0 & \text{if } x \notin E \end{cases}$$

Then for any event F , we have

$$P(F|E) = \int_F f(x|E)dx$$