The Scientist Game

Scott S. Emerson, M.D., Ph.D.

Professor of Biostatistics
University of Washington

1

Overview

- A simplified universe
  - One dimensional universe observed over time
  - Each position in the universe has an object
  - Goal is to discover any rules that might determine which objects are in a given location at a particular time

2

Objects

- Objects in the universe have only three characteristics, each with only two levels
  - Color: White or Orange
  - Size: BIG or small
  - Letter: A or B

A a B b A a B b

3

Universal Laws

- The level of each characteristic (color, size, letter) for the object at any position in the universe is either
  - completely determined by the prior sequence of that characteristic for objects at that position,

    OR
  - is completely random (anything is permissible)
- (No patterns involving probabilities less than 1)
- (Adjacent positions have no effect)

4

Universal Laws

- Furthermore any pattern to the objects at a position over time is “stationary”
  - The exact pattern repeats itself over a finite period of time

5

Examples of Universal Laws

- Color only (a cycle of length 2):

b a A b a a a B A a A A

- The next object in the sequence must be white, but any size or letter will do:

a A b B

6

Examples of Universal Laws

- Size and letter (a cycle of length 4):

A a B b A a B b A a B b

- The next object in the sequence must be a big A, but any color will do

A A

7

Examples of Universal Laws

- Size only (a cycle of length 2):

B a B b A b B a B b A a

- The next object in the sequence must be big, but any color or letter will do:

A B A B

8

Examples of Universal Laws

- No discernible pattern (in available data):

A a b a A b B a B B A a

- If there is truly no deterministic pattern, then any object may appear next:

a A b B a A b B

9

Scientific Task

- Goal is therefore to decide for some position
  - whether a rule governs the level of each characteristic, and
  - if so, what that rule is (pattern to the sequence)

10

Hypothesis Generation

- Initially we have observational data gathered over time
  - Amount of available information varies from position to position
  - We want to identify some position that is the most likely to be governed by some deterministic rule

11

Observational Data

	Time												
Pstn	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2
…														
114.	b	a	A	b	a	a	B	a	A	A	a	b	?	?
115.	A	a	B	b	A	a	b	b	A	a	a	A	?	?
116.							B	b	A	a	a	b	?	?
117.						b	B	A	b	b	b	a	?	?
118.	A	b	B	A	B	b	A	B	B	a	B	B	?	?
119.		B	b	B	b	A	a	A	a	B	A	b	?	?
120.										B	B	b	?	?
121.				B	A	a	B	b	b	a	b	A	?	?
…														

12

Observational Data

Time

```
Pstn -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1  0  1  2
…

118.   A   b  B  A  B  b  A  B  B  a  B  B  ?  ?

…
```

13

Next Step?

- Further observation?
  - Might take too long
  - Won't really establish cause and effect

14

Experimentation

- You can try to put an object in the position
  - If it cannot come next, it disintegrates and you can try another
  - If it can come next, it stays and you can try a different object to follow it

15

Experimental Goal

- You need to devise a series of experiments to discover
  - whether a deterministic rule governs the sequence of objects at position 118, and
  - if there is such a rule, what it is

16

Real World

- Problem:
  - You must buy objects to experiment with
    - (apply for a grant)
- Question:
  - What object should you try next in the sequence in order to determine the rule?

17

Possible Experiments

	Time												
Pstn	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2
118.	A	b	B	A	B	b	A	B	B	a	B	B	?	?

Possible Experiments

a A b B a A b B

- Which experiment do you do first?

18

Results of Observation

	Time												
Pstn	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2
118.	A	b	B	A	B	b	A	B	B	a	B	B	?	?

- We identified position 118 which had some regular patterns
  - Color cycle of length 2: (orange, white)
  - Size cycle of length 4: (big, little, big, big)
  - Letter cycle of length 3: (A, B, B)

19

Define Hypotheses

- Deterministic pattern versus random chance for each characteristic
  - Recognize that some or all observed patterns might be coincidence
    - Chance observation of a pattern for a single characteristic (e.g., color) with sample size 12 (assuming each level equally likely)
      - 1 out of 1024 for a cycle of length 2
      - 1 out of 512 for a cycle of length 3
      - 1 out of 256 for a cycle of length 4

20

Possible Hypotheses

- Assuming sufficient data to see any rule

118. A b B A B b A B B a B B ? ?

Hypotheses
Color, Size, and Letter
Color, Size
Color, Letter
Size, Letter
Color
Size
Letter
All coincidence

21

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Possible Experiments

Hypotheses | A
Color, Size, and Letter
Color, Size
Color, Letter
Size, Letter
Color
Size
Letter
All coincidence

22

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Possible Experiments

Hypotheses | A
Color, Size, and Letter | +
Color, Size
Color, Letter
Size, Letter
Color
Size
Letter
All coincidence

23

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Possible Experiments

Hypotheses | A
Color, Size, and Letter | +
Color, Size | +
Color, Letter
Size, Letter
Color
Size
Letter
All coincidence

24

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments A
Color, Size, and Letter	+
Color, Size	+
Color, Letter	+
Size, Letter	
Color	
Size	
Letter	
All coincidence	

25

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments A
Color, Size, and Letter	+
Color, Size	+
Color, Letter	+
Size, Letter	+
Color	
Size	
Letter	
All coincidence	

26

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments A
Color, Size, and Letter	+
Color, Size	+
Color, Letter	+
Size, Letter	+
Color	+
Size	
Letter	
All coincidence	

27

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments A
Color, Size, and Letter	+
Color, Size	+
Color, Letter	+
Size, Letter	+
Color	+
Size	+
Letter	
All coincidence	

28

Most Popular First Choice

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments: A
Color, Size, and Letter	+
Color, Size	+
Color, Letter	+
Size, Letter	+
Color	+
Size	+
Letter	+
All coincidence	

29

Most Popular First Choice

- A noninformative experiment

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments: A
Color, Size, and Letter	+
Color, Size	+
Color, Letter	+
Size, Letter	+
Color	+
Size	+
Letter	+
All coincidence	+

30

Next Worse Choice

- If all hypotheses equally likely, a 7-1 split

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments: A	b
Color, Size, and Letter	+	–
Color, Size	+	–
Color, Letter	+	–
Size, Letter	+	–
Color	+	–
Size	+	–
Letter	+	–
All coincidence	+	+

31

Other Suboptimal Experiments

- If all hypotheses equally likely, a 6-2 split

118. A b B A B b A B B a B B ? ?

Hypotheses	Possible Experiments: A	b	b	B	a
Color, Size, and Letter	+	–	–	–	–
Color, Size	+	–	–	–	–
Color, Letter	+	–	–	–	–
Size, Letter	+	–	–	–	–
Color	+	–	+	–	–
Size	+	–	–	+	–
Letter	+	–	–	–	+
All coincidence	+	+	+	+	+

32

Optimal Experiments

- Based on a binary search

118. A b B A B b A B B a B B ? ?

	Possible Experiments							
Hypotheses	A	b	b	B	a	B	a	A
Color, Size, and Letter	+	–	–	–	–	–	–	–
Color, Size	+	–	–	–	–	+	–	–
Color, Letter	+	–	–	–	–	–	+	–
Size, Letter	+	–	–	–	–	–	–	+
Color	+	–	+	–	–	+	+	–
Size	+	–	–	+	–	+	–	+
Letter	+	–	–	–	+	–	+	+
All coincidence	+	+	+	+	+	+	+	+

33

What If Data Insufficient?

- Suppose deterministic cycle length > 12

118. A b B A B b A B B a B B ? ?

	Possible Experiments							
Hypotheses	A	b	b	B	a	B	a	A
Color, Size, and Letter	+	–	–	–	–	–	–	–
Color, Size	+	–	–	–	–	+	–	–
Color, Letter	+	–	–	–	–	–	+	–
Size, Letter	+	–	–	–	–	–	–	+
Color	+	–	+	–	–	+	+	–
Size	+	–	–	+	–	+	–	+
Letter	+	–	–	–	+	–	+	+
All coincidence	+	+	+	+	+	+	+	+
Cycle length > 12	?	?	?	?	?	?	?	?

34

Moral

- The goal of the experiment should be to "decide which" not "prove that"
- A well designed experiment discriminates between hypotheses
  - The hypotheses should be the most important, viable hypotheses

35

Moral

- All other things being equal, an experiment should be equally informative for all possible outcomes
  - In the presence of a binary outcome, use a binary search
    - (using prior probability of being true)
  - But may need to consider simplicity of experiments, time, cost

36

In the Presence of Variability

- We use statistics to quantify the precision of our inference
  - We will describe our confidence/belief in our conclusions using frequentist or Bayesian probability statements

37

Statistical Experimental Design

- I believe a scientific approach to the use of statistics is to
  - Decide a level of confidence used to construct frequentist confidence intervals or Bayesian credible intervals
  - Ensure adequate statistical precision (sample size) to discriminate between relevant scientific hypotheses
    - The intervals should not contain two hypotheses that were to be discriminated between

38

Impact on Statistical Power

- I choose equal one-sided type I and type II errors
  - E.g., 97.5% power to detect the alternative in a one-sided level 0.025 hypothesis test
- In this way, at the end of the study, the 95% CI will not contain both the null and alternative hypotheses
  - I will have discriminated between the hypotheses with high confidence

39