Lesson 15

Tuesday 3/26/24

A Judge Hands Out 10 Prison Sentences

data = 7,2,1,3,8,4,5,5,3,7

<u>Given</u> these 10 cases, what is the probability that a case drawn at random is sentenced to more than 5 years in prison?

Which cases in this dataset received a prison sentence exceeding 5 years?

data = <u>7</u>,2,1,3,<u>8</u>,4,5,5,3,<u>7</u>

$$p(\text{sentence} > 5 \text{ years}) = \frac{3}{10} = 0.3$$

Seven people are released from prison. Each of these 7 people are followed for 3 years. At the end of the 3 years, we can classify each person as a success (S) or a failure (F) in terms of recidivism. Here is the data:

data = S,F,F,S,F,S,F

If we assume these people are independent of each other (like coin flips are independent), we can use this information to calculate the probability that someone selected at random from these data is observed to fail:

$$p(\text{fail}) = \frac{4}{7} = 0.571$$

With p(fail) in hand (0.571), we are now asked to calculate the probability distribution of the number of failures for 5 new prison releasees.

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	
2	
3	
4	
5	

k = 0

$$n! = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

$$k! = 0! = 1$$

$$(n-k)! = (5-0)! = 5! = 120$$

$$p^k = 0.571^0 = 1$$

$$(1-p)^{n-k} = (1-0.571)^{5-0} = 0.429^5 = 0.015$$

$$\frac{120}{1 \times 120} \times 1 \times 0.015 = 0.015$$

k = 1

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	0.097
2	
3	
4	
5	

$$n! = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

$$k! = 1! = 1$$

$$(n-k)! = (5-1)! = 4! = 24$$

$$p^k = 0.571^1 = 0.571$$

$$(1-p)^{n-k} = (1-0.571)^{5-1} = 0.429^4 = 0.034$$

$$\frac{120}{1 \times 24} \times 0.571 \times 0.034 = 0.097$$

k = 2

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	0.097
2	0.258
3	
4	
5	

$$n! = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

$$k! = 2! = 2$$

$$(n-k)! = (5-2)! = 3! = 6$$

$$p^k = 0.571^2 = 0.326$$

$$(1-p)^{n-k} = (1-0.571)^{5-2} = 0.429^3 = 0.079$$

$$\frac{120}{2 \times 6} \times 0.326 \times 0.079 = 0.258$$

k = 3

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	0.097
2	0.258
3	0.342
4	
5	

$$n! = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

$$k! = 3! = 6$$

$$(n-k)! = (5-3)! = 2! = 2$$

$$p^k = 0.571^3 = 0.186$$

$$(1-p)^{n-k} = (1-0.571)^{5-3} = 0.429^2 = 0.184$$

$$\frac{120}{6 \times 2} \times 0.186 \times 0.184 = 0.342$$

k = 4

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	0.097
2	0.258
3	0.342
4	0.227
5	

$$n! = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

$$k! = 4! = 24$$

$$(n-k)! = (5-4)! = 1! = 1$$

$$p^k = 0.571^4 = 0.106$$

$$(1-p)^{n-k} = (1-0.571)^{5-4} = 0.429^1 = 0.429$$

$$\frac{120}{24 \times 1} \times 0.106 \times 0.429 = 0.227$$

k = 5

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	0.097
2	0.258
3	0.342
4	0.227
5	0.061

$$n! = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

$$k! = 5! = 120$$

$$(n-k)! = (5-5)! = 0! = 1$$

$$p^k = 0.571^5 = 0.061$$

$$(1-p)^{n-k} = (1-0.571)^{5-5} = 0.429^0 = 1$$

$$\frac{120}{120 \times 1} \times 0.061 \times 1 = 0.061$$

With p(fail) in hand (0.571), we are now asked to calculate the probability distribution of the number of failures for 5 new prison releasees.

$$p(k \text{ failures out of 5 releasees}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}, k = 0, 1, ..., 5$$

k	p(k)
0	0.015
1	0.097
2	0.258
3	0.342
4	0.227
5	0.061

Now that we have this distribution, we are in a position to estimate quantities like:

p(3 or 4 or 5 people fail) =
p(3 or more people fail) =

0.342 + 0.227 + 0.061 = 0.63

or, p(no failures) = 0.015

Hypothesis Testing

**In lay terms, an hypothesis is an idea, explanation, or statement that can be tested; truth value of the statement can be assessed with empirical evidence.

Hypothesis Testing

In classical probability and statistics, we think of hypotheses in terms of the value or magnitude of a scientifically interesting parameter (i.e., a mean, a proportion, a correlation, etc.)

Deductive Inference

1. State hypothesis

2. Collect appropriate evidence

3. Use a valid probability calculation to test the hypothesis

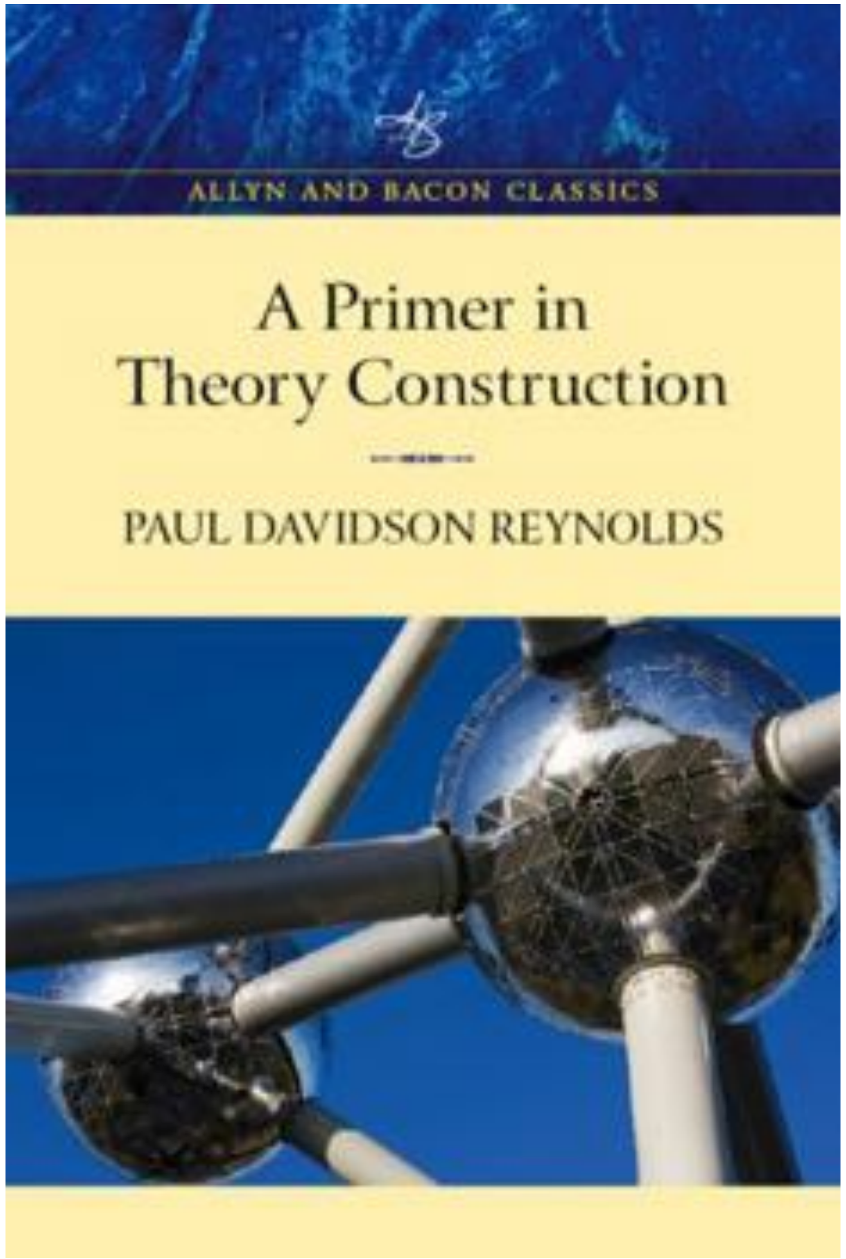
4. Draw conclusion

5. Go back to #1

(Scientific Method)

Inductive Inference

Reynolds (1977) distinguishes between deductive and inductive inference: (1) deductive means that the hypothesis is stated before data are collected; (2) inductive means that data are collected and then a hypothesis is developed based on the evidence. Both are valid and necessary!



It might have occurred to you that we could think about...

$$p(\text{H is correct}|\text{evidence})$$

where *H* stands for the term "Hypothesis" or "Model"; this term might look familiar as it is similar to a term we considered alongside the general multiplication rule.

What is p(H at Time 1 and H at Time 2)?

Reminder!

General multiplication rule:

$$p(H_1 \; \& \; H_2) = p(H_1) \times p(H_2 | H_1)$$

Does Not Assume Independence

flip a coin - time 2

flip a coin - time 1

	Heads	Tails	Total
Heads	246	250	496
Tails	251	253	504
Total	497	503	1000

Multiplicative Parts:

p(H at Time 1) = $p(H_1)$ = 496/1000
$p(H_2|H_1)$ = 246/496 = 246/496

Solution:

496/1000 × 246/496 =

0.496 × 0.496 = 0.246

compare to what we got on previous slide -- evidence that the two events are independent.

How would this work?

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)}$$

Rev. Thomas Bayes
(1701-1761)

Conditional probability rule derived from Bayes' Theorem; you can read about Thomas Bayes at his wikipedia page (link).

Laplace was among the first to recognize conditional probability as a framework for inference.

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)}$$

Pierre-Simon Laplace (1749-1827)

Laplace famously considered the probability that the sun will rise tomorrow (so H = the sun will rise tomorrow -- an idea that can be tested). He concluded that we should not approach such a question de novo; rather we should consider existing knowledge and understanding of historical daily sunrise patterns to evaluate this probability (we are given the evidence and asked to calculate the probability).

Wikipedia - sunrise problem article (link).

Scholars like Ronald Fisher had concerns...

?

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)}$$

Ronald Fisher (1880-1962)

Instead, we should evaluate:

1

$$p(E|H) = \frac{p(H|E)p(E)}{p(H)}$$

1

Condition on the truth of a hypothesis and assess the probability of the data looking the way they do.

Page 169 of your book:

"Recall that we begin our hypothesis test by assuming that the null hypothesis is true..."

"What we do, then, is maintain our belief in the null hypothesis until we are informed by the data that this belief is improbable given what we have observed." [Emphasis added]

For now, let's ignore the "null" adjective and just consider the word "hypothesis."

Let's consider a real study...

CRIME AND THE BUSINESS CYCLE

*PHILIP J. COOK and GARY A. ZARKIN**

THE business cycle has a pervasive effect on the structure of economic opportunity and hence on behavior. The effect is reflected in social indicators as diverse as school enrollments, birthrates, and labor force participation.[1] It would be surprising indeed if crime rates were immune to general business conditions, and certainly the conventional wisdom asserts that ''street'' crime is countercyclical. A recession always provides police chiefs with a comfortable explanation for their failure to prevent increases in the crime rate.

Question: when the economy contracts, what happens to crime?

Let's think about robbery.

From the 1930's to the early 1980's, there were 9 complete business cycles (NBER Business Cycle Dating Committee)

August 1929 (1929Q3)	March 1933 (1933Q1)	43	21	64	34
May 1937 (1937Q2)	June 1938 (1938Q2)	13	50	63	93
February 1945 (1945Q1)	October 1945 (1945Q4)	8	80	88	93
November 1948 (1948Q4)	October 1949 (1949Q4)	11	37	48	45
July 1953 (1953Q2)	May 1954 (1954Q2)	10	45	55	56
August 1957 (1957Q3)	April 1958 (1958Q2)	8	39	47	49
April 1960 (1960Q2)	February 1961 (1961Q1)	10	24	34	32
December 1969 (1969Q4)	November 1970 (1970Q4)	11	106	117	116
November 1973 (1973Q4)	March 1975 (1975Q1)	16	36	52	47
January 1980 (1980Q1)	July 1980 (1980Q3)	6	58	64	74
July 1981 (1981Q3)	November 1982 (1982Q4)	16	12	28	18

When each business cycle ends, the economy transitions from a growth phase to a contraction phase (a recession).

For each of the 9 times the economy tipped into a recession, C&Z asked the question, "Did robbery rates move higher relative to their yearly trends during the growth phase?"

Notice this is a yes or no question.

Suppose we assume that there is no relationship between changes in the economy and changes in crime.

What would we expect to see?

We might expect to see that when the economy tips into recession, robberies would be equally likely to increase or decrease (relative to their trend in the growth phase). Like a coin flip would be equally likely to come up heads or tails.

This a hypothesis -- what your book would call a null hypothesis.

C&Z collected their data and found that in 8 out of the 9 business cycles that robbery increased relative to its growth phase trend when the economy tipped into recession.

In other words, they found what they called a strong countercyclical pattern.

The question is whether the evidence is strong enough to reject the hypothesis of equally likely outcomes.

What is the probability distribution for the number of "countercyclical movements"?

--

assuming that a growth or decline in robberies is equally likely when the economy moves into a recession.

$$p(k \text{ countercyclical movements}|p = 0.5, N = 9) = \frac{N!}{k!(N-k)!}p^k(1-p)^{N-k}$$

$$p(k \text{ countercyclical movements} | p = 0.5, N = 9) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k}$$

k	p(k)
0	0.002
1	0.018
2	0.070
3	0.164
4	0.246
5	0.246
6	0.164
7	0.070
8	0.018
9	0.002

p(0 or 1 CC Movements) = 0.002 + 0.018 = 0.020

p(2-7 CC Movements) = 0.070+0.164+0.246+0.246+0.164+0.070 = 0.96

p(8 or 9 CC Movements) = 0.018 + 0.002 = 0.020

What have we learned? Getting 0 or 1 or 8 or 9 CC movements is a rare occurrence <u>if</u> p = 0.5...