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 (a set of Bernoulli trials). 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

Unconditional vs. Conditional Probability

Suppose we observe a sample of 3200 people who are 25 years old. For each of these people, we check their criminal history records and we determine that 448 of them had attained at least one criminal conviction after turning 18 years old. Then,

$$p(\text{Convicted at least once}) = \frac{\#\text{ of times an event occurred}}{\#\text{ of times it could have occurred}} = \frac{448}{3200} = 0.14$$

Now, suppose we ask a slightly different question: among the 679 people who were arrested at least one time after turning 18, what is the probability of at least one conviction?

p(Convicted at least once | Arrested) = 448/679 = 0.660

Let's Consider the Same Two Calculations in Table Form

	Arrest= No	Arrest= Yes	Total
Convict = No	2521	231	2752
Convict = Yes	0	448	448
Total	2521	679	3200

p(convict) = 448/3200 = 0.140

p(arrest|convict) = 448/448 = 1

p(arrest) = 679/3200 = 0.212

p(Convict|Arrest) = p(convict) x p(Arrest|Convict) / p(arrest)

0.140 1 0.212

Reduces to: p(convict|arrest) = 0.140/0.212 = 0.660

***Very important to pay attention to the conditioning.

Hypothesis Testing

**In lay terms, a 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.).

Statistical Inference

Based on the evidence we have, a statement is made about a scientifically interesting but unobserved parameter (an extrapolation). Inferences are comprised of two components: (1) an estimate; and (2) a measure of uncertainty to accompany the estimate.

Deductive Science

1. State hypothesis

2. Collect appropriate evidence

3. Use a valid probability calculation to test the hypothesis

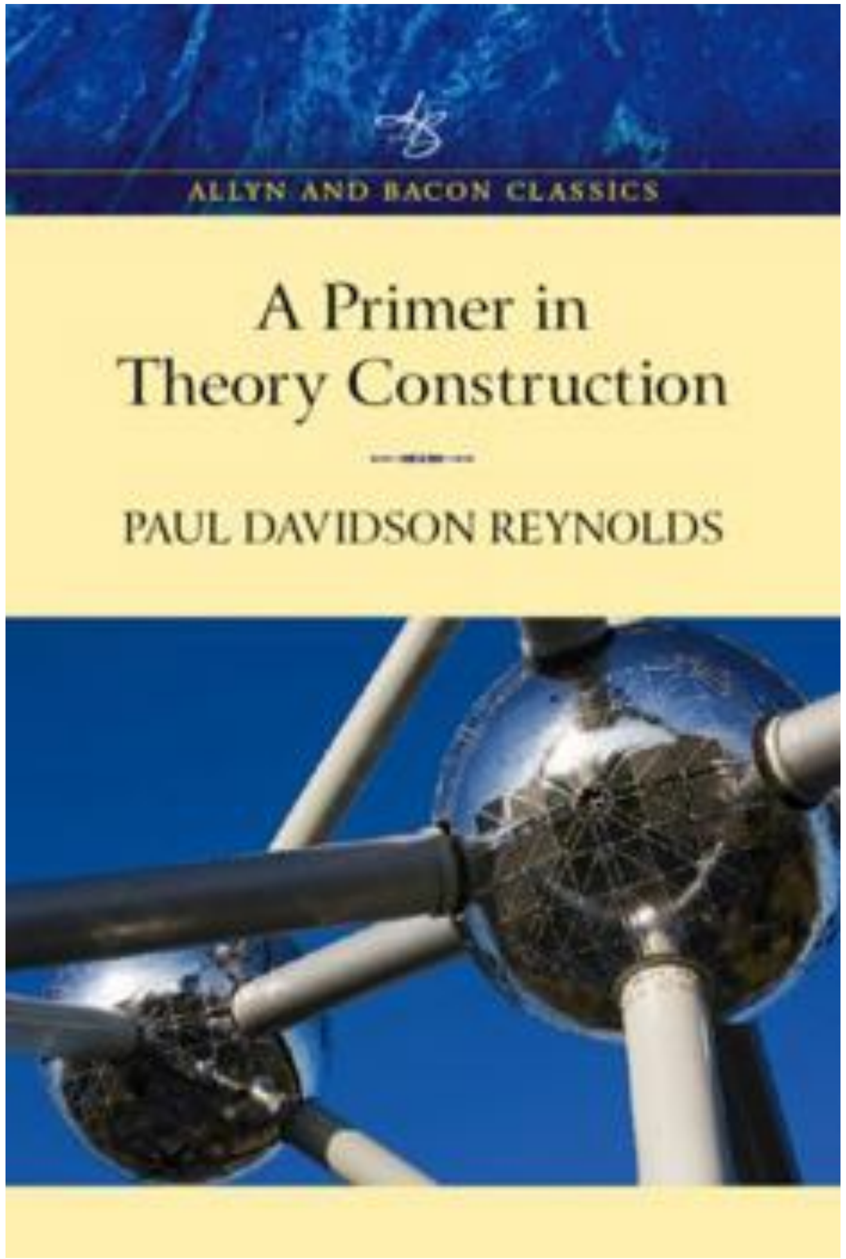
4. Draw conclusion

5. Go back to #1

(Scientific Method)

Inductive Science

Reynolds (1977) distinguishes between deductive and inductive science: (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!



Hypotheses and Evidence

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

"Inverse Probability" or Bayesian Inference

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

Classical Inference

At top of page 167, your book says "In a hypothesis test, the null hypothesis is the hypothesis that is intially *assumed to be true.*" Thus, classical inference focuses on the probability of the evidence looking the way it does conditional on the truth of a particular hypothesis.