Lesson 14

Thursday 3/14/24

2 Coin Flip Experiment

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

There are a few probability-related questions we could ask about this.

Exhaustive Set of Outcomes - Flip a coin once.

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

Exhaustive: if we flip a coin one time, there are 2 and only 2 possible outcomes. Remember mutual exclusivity means that only one of the outcomes is possible on a single coin flip. Exhaustive means that H and T are the only possible outcomes. These 2 outcomes exhaust all of the possibilities.

S = {H,T}

S means "sample space"

Exhaustive Set of Outcomes - Flip a coin twice

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

Exhaustive: flip a coin twice. What is the set of possible outcomes?

$$S = \{HH, HT, TH, TT\}$$

There are no possibilities besides these 4. In other words, these 4 outcomes exhaust all of the possibilities.

Mutual exclusivity - flip a coin once

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

Mutual exclusivity: the outcome of a single coin flip must be H or T. It cannot be both.

$$S = \{H, T\}$$

Mutual exclusivity - flip a coin twice

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

Mutual exclusivity: do the coin flip experiment twice. What is the set of mutually exclusive outcomes?

$$S = \{HH, HT, TH, TT\}$$

If you flip a coin twice, the outcome can be one and only one of the elements in this set. It can't be more than one.

What is p(H at Time 1 <u>or</u> H at Time 2)?

Restricted addition rule:

$$p(H_1+H_2) = p(H_1) + p(H_2)$$

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

Additive Parts:

p(H at Time 1) = 496/1000
p(H at Time 2) = 497/1000

Solution:

496/1000 + 497/1000 =

0.496 + 0.497 = 0.993

Notice that the 246 is being counted twice here...

What is p(H at Time 1 or H at Time 2)? (continued)

General addition rule:

$$p(H_1+H_2) = p(H_1) + p(H_2) - p(H_1 \text{ and } H_2)$$

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

Additive Parts:

p(H at Time 1) = 496/1000
p(H at Time 2) = 497/1000
p(H at Time 1 & H at Time 2) = 246/1000

Solution:

496/1000 + 497/1000 - 246/1000 =

0.496 + 0.497 - 0.246 = 0.747

Independence of 2 Variables - each coin flip is a variable

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

Independence: suppose I flip a fair coin 2 times. If I know the outcome of the first flip, does that help me predict the outcome of the second flip?

If the answer is "no", then the 2 flips are independent of each other.

So, how can we tell? We could use our common sense but there is also a way we can check...

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

Restricted multiplication rule:

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

Assumes 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 at Time 2) = $p(H_1)$ = 497/1000

Solution:

$496/1000 \times 497/1000 =$

$0.496 \times 0.497 = 0.247$

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

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

<u>Multiplicative Parts</u>:

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

<u>Solution</u>:

496/1000 × 246/496 =

0.496 × 0.496 = 0.246

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

What is p(carry a weapon or commit a crime)?

	Weapon = No	Weapon = Yes	Total
Crime = No	335	11	346
Crime = Yes	70	16	86
Total	405	27	432

Restricted Addition Rule

$$p(a+b) = p(a) + p(b)$$

$$p(w+c) = p(w) + p(c) = 27/432 + 86/432$$

$$= 0.063 + 0.199 = 0.262$$

**problem is that the 16 people are being counted twice.*

What is p(carry a weapon or commit a crime)?

(Continued)

	Weapon = No	Weapon = Yes	Total
Crime = No	335	11	346
Crime = Yes	70	16	86
Total	405	27	432

General Addition Rule

$$p(a+b) = p(a) + p(b) - p(a\&b)$$

$p(w+c) = p(w) + p(c) - p(w \& c) =$

$27/432 + 86/432 - 16/432 =$

$= 0.063 + 0.199 - 0.037 = 0.225$

What is p(carry a weapon and commit a crime)?

	Weapon = No	Weapon = Yes	Total
Crime = No	335	11	346
Crime = Yes	70	16	86
Total	405	27	432

Restricted Multiplication Rule - Assumes Independence

$$p(a \ \& \ b) = p(a) \times p(b)$$

$$p(w \ \& \ c) = p(w) \times p(c) = 27/432 \times 86/432$$

$$= 0.063 \times 0.199 = 0.013$$

What is p(carry a weapon <u>and</u> commit a crime)?

(Continued)

	Weapon = No	Weapon = Yes	Total
Crime = No	335	11	346
Crime = Yes	70	16	86
Total	405	27	432

General Multiplication Rule - Does Not Assume Independence

$$p(a \ \& \ b) = p(a) \times p(b|a)$$

$$p(w \ \& \ c) = p(w) \times p(c|w) = 27/432 \times 16/27$$

$$= 0.063 \times 0.593 = 0.037$$

Suppose previous research tells us that about 40% of juvenile probationers will fail a drug test. The local juvenile court gives a drug test to the first 7 juvenile probationers this month. What is the probability distribution for the variable that measures the following quantity: "out of the 7 drug tests administered, how many will be positive"?

$S_{BT} = \{0,1\}$ → Sample space for each individual drug test (0=neg, 1=pos)

$S_C = \{0,1,2,3,4,5,6,7\}$ → Sample space for number of positive tests when 7 individual persons are tested.

**In both sample spaces, the elements are mutually exclusive and exhaustive; note that S_{BT} stands for the sample space of a Bernoulli trial (variable); S_C stands for the sample space of a counted outcome variable.

Suppose previous research tells us that about 40% of juvenile probationers will fail a drug test. The local juvenile court gives a drug test to the first 7 juvenile probationers this month. What is the probability distribution for the variable that measures the following quantity: "out of the 7 drug tests administered, how many will be positive"?

Let's start by estimating the probability that none of the 7 tests is positive:

$$p(0 \,\&\, 0 \,\&\, 0 \,\&\, 0 \,\&\, 0 \,\&\, 0 \,\&\, 0) =$$

$$0.6 \times 0.6 \times 0.6 \times 0.6 \times 0.6 \times 0.6 \times 0.6 = 0.028$$

**Assumption: each of the 7 tests are independent of each other; we can use restricted multiplication rule.

**Note: Each of the 7 drug tests is considered an independent Bernoulli trial; the probability of zero positive drug tests is governed by a Bernoulli process.

The calculation on the previous slide is do-able but it is tedious. Another way might be easier.

$$p(0 \text{ positive drug tests}) = \binom{7}{0} \times 0.4^0 \times (1 - 0.4)^{7-0}$$

$$=$$

$$p(0 \text{ positive drug tests}) = \frac{7!}{0!(7-0)!} \times 0.4^0 \times (1 - 0.4)^{7-0}$$

Convince yourself that you will get the same answer as the previous slide...

Let's try the next one now: what is the probability of exactly one positive drug test, out of the 7 administered?

$$p(1 \text{ positive drug test}) = \frac{7!}{1!(7-1)!} \times 0.4^1 \times (1-0.4)^{7-1}$$

$$p(1 \text{ positive drug test}) = \frac{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{1 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1} \times 0.4^1 \times 0.6^6$$

$$p(1 \text{ positive drug test}) = \frac{5040}{720} \times 0.4^1 \times 0.6^6$$

$$p(1 \text{ positive drug test}) = 7 \times 0.4 \times 0.047$$

Solution: 0.132

We are now ready to write the *binomial probability mass function* for this problem:

$$p(\text{k positive drug tests out of n tests}) = \frac{n!}{k!(n-k)!} \times p^k \times (1-p)^{n-k}$$

for k = 0, 1, 2, ..., 7 and p = 0.4 (given)

k	p(k)
0	0.028
1	0.131
2	0.262
3	0.290
4	0.194
5	0.077
6	0.017
7	0.002

Verify: $\sum_{k=0}^{7} p(k) = 1.0$

We can now answer questions like "what is our best guess for the number of positive drug tests in a sample of 7 people?" (assuming that the previously understood probability of a positive test is 0.4.