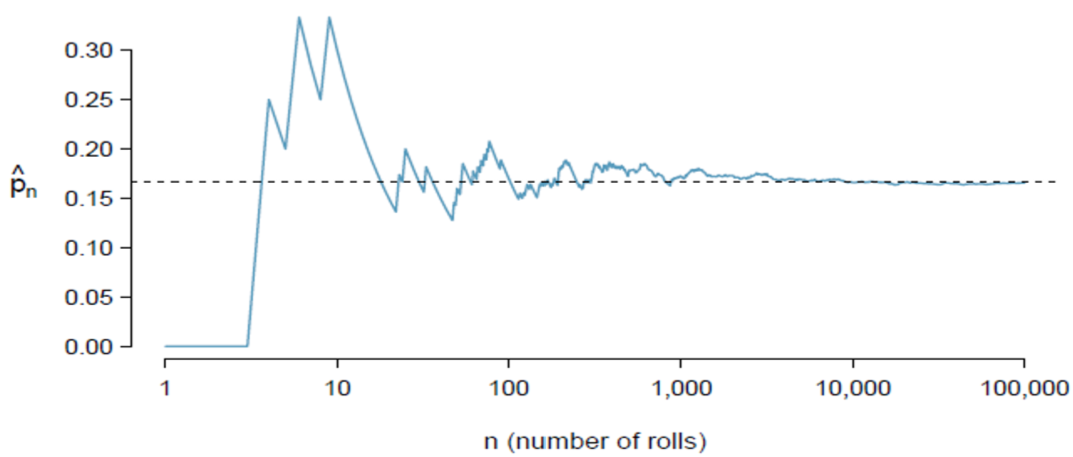
Topic 8: Basics of Probability

First Examples and Terminology

A **random phenomenon** is ...

Examples of random phenomena

Example: Rolling a 6-sided die:

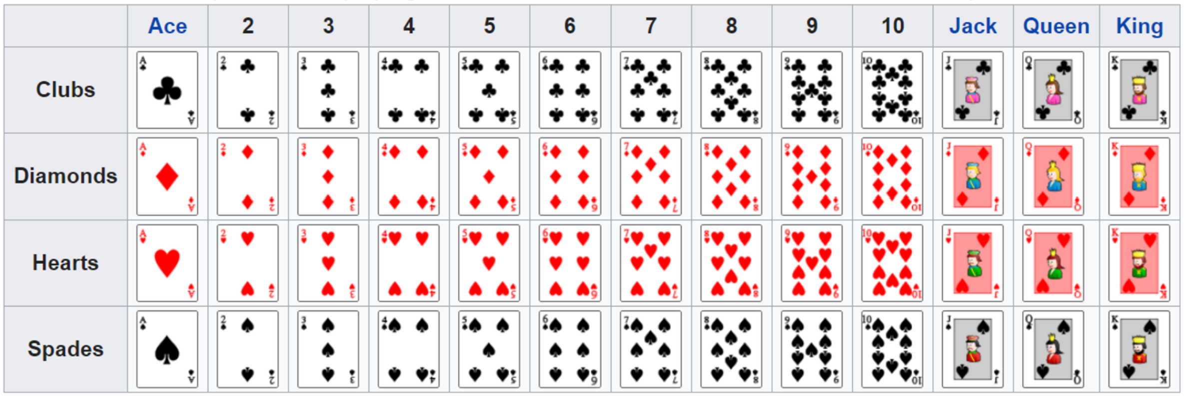


Let $\hat{p}_n =$

Let $p =$

The tendency for $\hat{p}_n$ to approach p as n increases is ...

Another example: Playing cards



Random phenomenon:

Event of interest:

Two ways to consider probability of this event:

- **Empirically**

- **Theoretically**

Building a **probability distribution**:

1. Describe random phenomenon

2. Describe **sample space**

3. Assign a **probability** to each outcome in the sample space

Random phenomenon: Flip a fair coin 3 times and note the outcomes (H or T) in order

- Sample space:

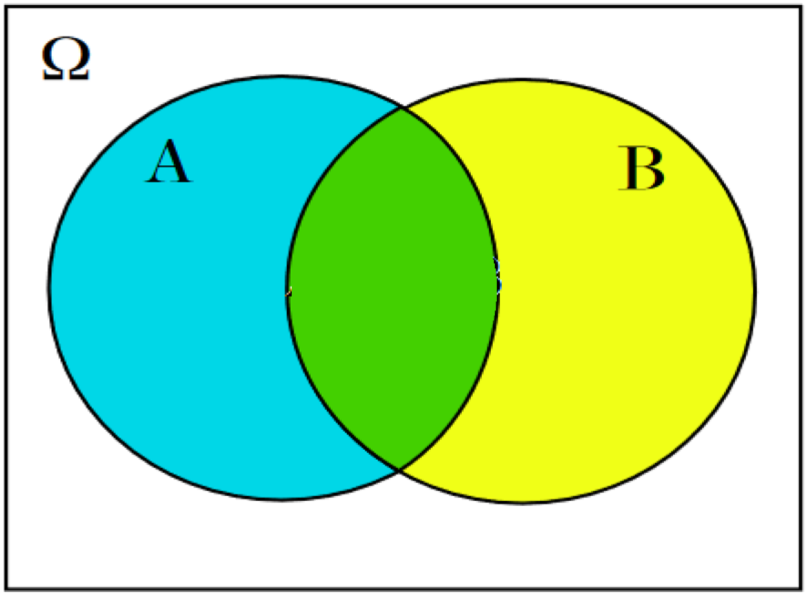
- Probabilities:

An **event** is ...

Back to coin flip example:

Probability of this event?

Venn diagram representation:



Event operators:

- **and**

- **or**

- **complement**

Random phenomenon: Toss a coin 10 times and record the number of times it lands on heads

Sample space?

Let's define some events:

Describe the event (A or B)

Describe the event (A and B)

Are A and C disjoint events?

Describe the event A^c

Four basic rules for event probabilities:

1. For any event A, $0 \leq P(A) \leq 1$.
2. $P(\Omega) = 1$.
3. If events A and B are disjoint, then $P(A \text{ or } B) = P(A) + P(B)$.
4. For any event A, $P(A^c) = 1 - P(A)$.

Example: Benford's Law

Random phenomenon: Take a table of financial, demographic, or similar numerical data, randomly select one observation, and check its first digit.

Sample space?

Benford's Law postulates that all digits are not equally likely!

Populations of NC Counties, 2023 Estimates

1	Wake	1,190,275	26	Robeson	117,365	51	Columbus	50,121	76	Madison	22,071
2	Mecklenburg	1,163,701	27	Moore	106,898	52	Duplin	49,520	77	Anson	21,897
3	Guilford	549,866	28	Craven	102,391	53	Edgecombe	48,832	78	Martin	21,447
4	Forsyth	392,921	29	Cleveland	101,378	54	Halifax	47,298	79	Greene	20,530
5	Cumberland	337,890	30	Nash	96,551	55	Stokes	45,532	80	Polk	20,060
6	Durham	336,892	31	Lincoln	95,675	56	McDowell	44,893	81	Hertford	19,453
7	Buncombe	275,901	32	Rockingham	92,518	57	Davie	44,599	82	Yancey	18,938
8	Union	256,452	33	Burke	88,338	58	Jackson	44,574	83	Warren	18,836
9	Johnston	241,955	34	Chatham	81,624	59	Beaufort	44,481	84	Avery	17,561
10	Cabarrus	240,016	35	Caldwell	80,574	60	Richmond	42,324	85	Bertie	16,922
11	New Hanover	238,852	36	Wilson	78,970	61	Vance	42,301	86	Northampton	16,715
12	Gaston	237,242	37	Franklin	77,001	62	Pasquotank	41,444	87	Mitchell	14,999
13	Onslow	213,676	38	Surry	71,462	63	Person	39,737	88	Swain	13,916
14	Iredell	199,710	39	Carteret	69,615	64	Macon	38,412	89	Chowan	13,891
15	Alamance	179,165	40	Pender	68,521	65	Dare	38,110	90	Perquimans	13,377
16	Pitt	175,119	41	Lee	67,059	66	Yadkin	37,774	91	Pamlico	12,423
17	Davidson	174,804	42	Wilkes	66,013	67	Alexander	36,473	92	Clay	11,864
18	Catawba	164,645	43	Stanly	65,699	68	Scotland	34,376	93	Alleghany	11,342
19	Brunswick	159,964	44	Rutherford	65,507	69	Transylvania	33,549	94	Camden	11,137
20	Rowan	151,661	45	Haywood	62,969	70	Currituck	31,593	95	Washington	10,713
21	Orange	150,626	46	Granville	62,192	71	Cherokee	29,959	96	Gates	10,343
22	Randolph	147,458	47	Sampson	59,601	72	Bladen	29,484	97	Jones	9,401
23	Harnett	141,477	48	Lenoir	54,895	73	Ashe	27,063	98	Graham	8,052
24	Henderson	119,230	49	Watauga	54,748	74	Montgomery	26,085	99	Hyde	4,607
25	Wayne	118,686	50	Hoke	54,446	75	Caswell	22,807	100	Tyrrell	3,461

1st digit	1	2	3	4	5	6	7	8	9
prop.	.34	.17	.12	.12	.06	.08	.03	.04	.04
B. Law	.301	.176	.125	.097	.079	.067	.058	.051	.046

Let X represent the first digit of a number picked at random from a data set that follows Benford's Law exactly. Calculate:

- $P(X = 1 \text{ or } X = 2)$
- $P(X > 2)$
- $P(X \text{ is even})$
- $P(X = 1 \text{ or } X \text{ is even})$

Random phenomenon: Roll three fair, six-sided dice and count the number of 6's that come up, represent this count with X

Sample space?

Probability distribution:

value of X				
probability	.579	.347	.069	.005

Calculate:

- $P(X = 2 \text{ or } X = 3)$

- $P(X \neq 0)$

Random phenomenon: Take a SRS of size 2 from a group of 5 people: Alice, Bob, Cindy, Dave, Eva

Sample space?

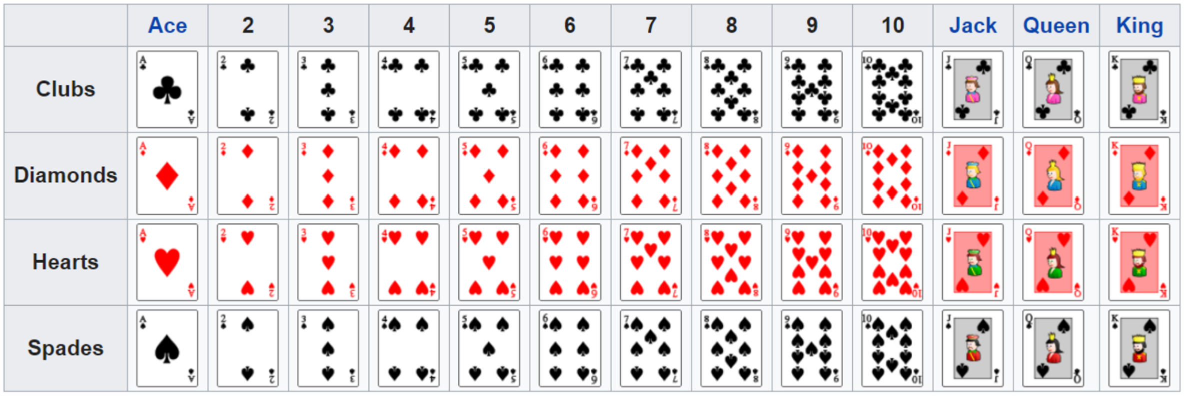
Probability distribution:

Calculate:

- $P(\text{Bob is chosen but Cindy is not})$

- $P(\text{Bob or Cindy is chosen})$

Random phenomenon: Draw the top card from a well-shuffled standard deck



Event A: The card drawn is a diamond.

Event B: The card drawn is a face card.

$P(A) =$

$P(B) =$

$P(A \text{ or } B) =$

A more general rule for calculating $P(A \text{ or } B)$:

Independence and the Multiplication Rule

Two events A and B are **independent** if ...

Examples:

Multiplication rule for independent events:

Random phenomenon: Roll a fair six-sided die twice, record both roll outcomes

Event A: The first roll comes up 6.

Event B: The second roll comes up 6.

Are A and B independent?

Are A and B disjoint?

$P(A \text{ and } B) =$

$P(A \text{ or } B) =$

Random phenomenon: Take SRS of size 2 from a group of 5 people: Alice, Bob, Cindy, Dave, Eva

Recall sample space: {Alice, Bob} {Alice, Cindy} {Alice, Dave} {Alice, Eva} {Bob, Cindy}
{Bob, Dave} {Bob, Eva} {Cindy, Dave} {Cindy, Eva} {Eva, Dave}

Event A: Alice is selected.

Event B: Bob is selected.

Are A and B independent?

Are A and B disjoint?

$P(A \text{ and } B) =$

Random phenomenon: Roll three fair, six-sided dice and count the number of 6's that come up, represent this count with X

Probability distribution:

value of X				
probability	.579	.347	.069	.005

We now know enough to compute two of these probabilities by hand!

Random phenomenon: From a very large population in which 40% identify as Democrats, take a SRS of size 5 and note whether or not each sampled person is a Democrat.

Event A: First person sampled is a Democrat

Event B: Second person sampled is a Democrat

Events C, D, and E defined similarly.

Are A, B, C, D, E independent?

$P(A \text{ and } B \text{ and } C \text{ and } D \text{ and } E) \approx$

Random phenomenon: Sample one American at random, check their hair color and eye color.

Event A: Person has blue eyes

Event B: Person has brown hair

If 27% of Americans have blue eyes and 53% of Americans have brown hair, what is $P(A \text{ and } B)$?