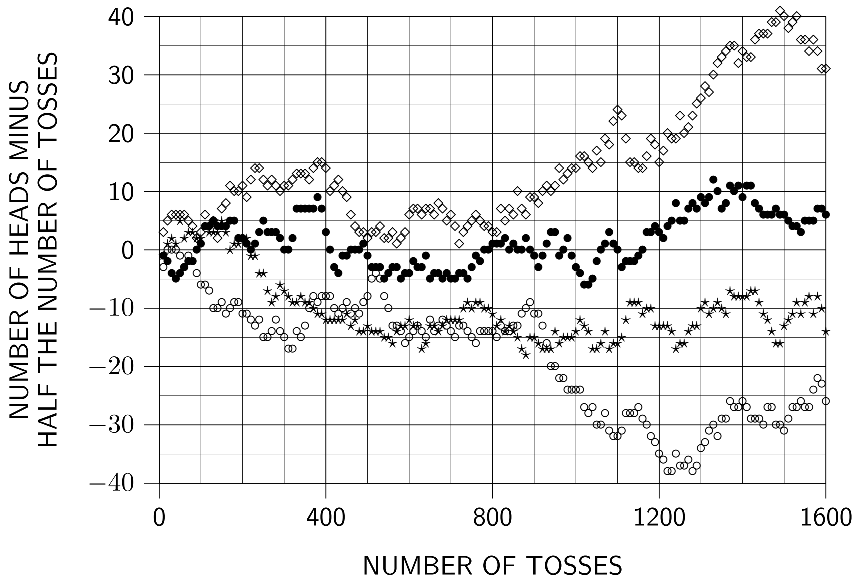
INTRODUCTION TO CHANCE VARIABILITY
WHAT DOES THE LAW OF AVERAGES SAY?

- 4 coins were tossed 1600 times each, and the “chance error”

$$\text{number of heads} - \text{half the number of tosses}$$

was plotted against the number of tosses:



- The plot suggests that if a coin is tossed 400 times, you’ll get ________ heads, give or take somewhere around ________ heads.
  - After 400 tosses, the chance errors were: $\diamond$ __, $\bullet$ 3, $\star$ −12, $\circ$ −8.
  - The overall size of these chance errors is measured by their RMS

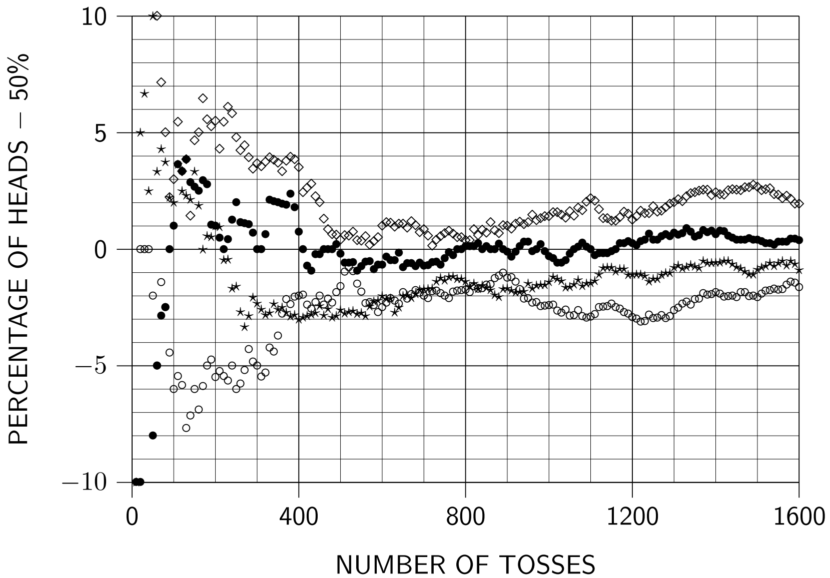
$$\sqrt{\left((\quad)^2 + 3^2 + (-12)^2 + (-8)^2\right)/4} = 10.1$$

- The plot also suggests that if a coin is tossed 1600 times, you’ll get ________ heads, give or take somewhere around ________ heads.
  - After 1600 tosses, the chance errors were: $\diamond$ __, $\bullet$ 6, $\star$ −14, and $\circ$ −26, with a RMS of

$$\sqrt{\left((\quad)^2 + 6^2 + (-14)^2 + (-26)^2\right)/4} = 21.6$$

- In general, as the number of tosses goes up, the difference between the number of heads and half the number of tosses tends to get ____________.
  - How big to you think the difference is likely to be for:
    - $400 \times 25 = 10{,}000$ tosses?
    - $10{,}000 \times 100 = 1{,}000{,}000$ tosses?

• For the same 4 coins, here are the chance errors expressed as a percentage of the number of tosses:



• If a coin is tossed 400 times, the percentage of heads will be ________%, give or take somewhere around ____________%.

- (200/400) × 100% = ________.
- (10/400) × 100% = ________.

• If a coin is tossed 1600 times, the percentage of heads will be ________%, give or take somewhere around ____________%.

- (20/1600) × 100% = ________.

• In general, as the number of tosses goes up, the difference between the percentage of heads and 50% tends to get ____________.

• To summarize, what does the so-called *law of averages* have to say about the

chance error = number of heads − half the number of tosses?

- The chance error is likely to be ________ in absolute terms, but ________ relative to the number of tosses.

• A coin lands heads 550 times in 1000 tosses. The chance error is:

- ____________, in absolute terms;
- ____________, as a percentage of the number of tosses.

• A coin lands heads 499,000 times in 1,000,000 tosses. The chance error is:

- ____________, in absolute terms;
- ____________, as a percentage of the number of tosses.

• A coin will be tossed, and you will win a dollar if the number of heads is within 5 of half the number of tosses. Which is better: 10 tosses or 100 tosses? Explain.

- ________ tosses is better.

• A coin will be tossed, and you will win a dollar if the number of heads is exactly equal to the number of tails. Which is better: 10 tosses or 100 tosses? Explain.

- ________ tosses is better.

• A coin will be tossed, and you will win a dollar if the percentage of heads is between 40% and 60%. Which is better: 10 tosses or 100 tosses? Explain.

- ________ tosses is better.

BOX MODELS

- Question: What do these things have in common?
  - The number of squares moved in a play at monopoly.
  - The number of heads in 10,000 tosses of a coin.
  - The amount on money won or lost gambling at roulette.
  - The number of Democrats in a random sample of voters.

Answer: All these numbers are subject to ______________________.

- Question: how can we analyze the chance variation?

Answer: Through a *box model*:

  - Set up an analogy between the chance quantity of interest and the sum of numbers drawn at random from some box.
  - Analyze the chance variability in the sum of the draws from the box.
  - Reformulate the results in terms of the quantity of interest, using the analogy.

THE SUM OF THE DRAWS

- A chance process for generating a number can often be modeled as follows:
  - There is a box of tickets.
  - Each ticket has a number written on it.
  - Some tickets are drawn at random from the box.
    - Typically with replacement — at each draw, all the tickets in the box have an equal chance of being drawn.
  - The numbers on the tickets are added up.
    - The total is called the *sum of the draws*.
- 100 draws are to be made from the box

| +1 | −2 |

  - The smallest the sum of the draws could be is ________.
  - The largest the sum of the draws could be is ________.
  - You'd expect [+1] to turn up about ________ times.
  - You'd expect [−2] to turn up about ________ times.
  - You'd expect the sum of the draws to be about ________.
- 100 draws are to be made with replacement from one of the following boxes. Your job is to guess what the sum of the draws will be, and you win $1 if you are right to within 10. In each case, what would you guess? Which box is best? Worst?

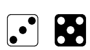
(i) | 1 9 | (ii) | 4 6 | (iii) | 5 5 |

  - Guesses: (i) ________, (ii) ________, (iii) ________.
  - Best box is ________. Worst box is ________.

- In making a box model, what three questions need to be answered?
  - What numbers go into the box?
  - How many of each kind?
  - How many draws?
- How do you answer the first two questions?
  - By thinking about what numbers the chance process generates, and the chances for each number.
- *Rolling a pair of dice.* The total number spots resulting from rolling a pair of dice is like the sum of ______ draws from the box

  | □ □ □ □ □ □ |

  - On a single roll of the die, any one of the numbers 1, 2, . . . , 6 could come up. So the tickets in the box should be labeled □, □, □, □, □, and □.
  - Each number is equally likely, so there should be ______ ticket of each kind.
  - Rolling a pair of dice is like making ______ draws from the box.
  - To illustrate the analogy:

Rolls of the dice:  Total number of spots: 8
Draws from the box: □ □ Sum of the draws: __

- *Tossing a coin.* The number of heads in 5 tosses of a fair coin is like the sum of ______ draws from the box

  | □ □ |

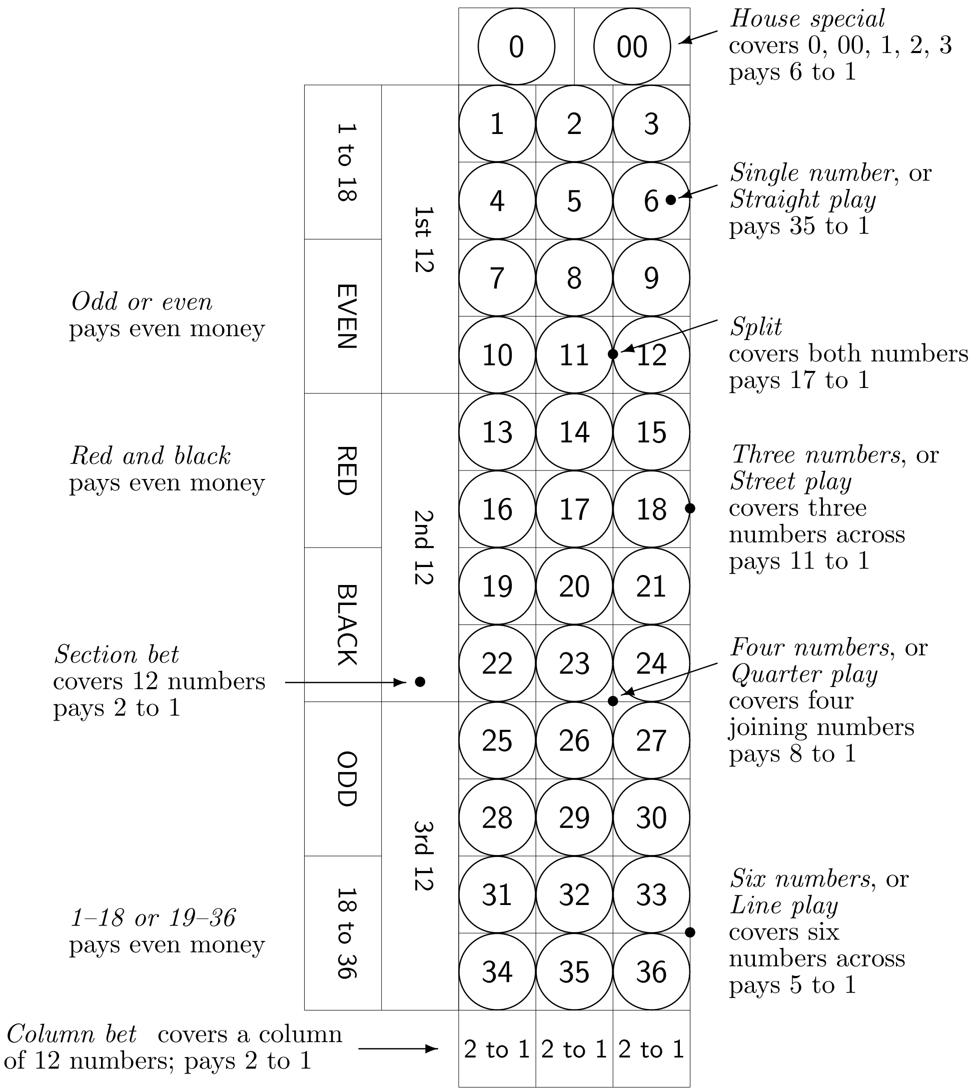
  where □ corresponds to getting a head, and □ corresponds to getting a tail.
  - Think about counting the number of heads in the 5 tosses. On each toss:
    - you add 1 to the count, if you get a head;
    - you add 0 to the count, if you get a tail.
  - The tickets in the box should be labeled □ and □.
  - Heads and tails are equally likely, so there should be ______ ticket of each kind.
  - There should be ______ draws.
  - To illustrate the analogy:

Tosses: H T H H T Number of heads: ________
Draws: □ □ □ □ □ Sum of the draws: ________

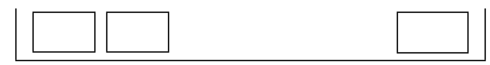
  - This is an example of "classifying and counting"; see section 5 of the next chapter.

A Nevada Roulette Table (FPPA, page 256)

Roulette is a pleasant, relaxed, and highly comfortable way to lose your money. — JIMMY THE GREEK



• *Playing roulette.* A gambler is going to play roulette 5 times, putting a dollar on 11 and 12 each time. If either of those numbers comes up, the gambler gets the dollar back, together with winnings of $17. If neither number comes up, he loses his dollar. The gambler's net gain in the 5 plays is like the sum of _______ draws from the box



• On each spin of the wheel, the gambler either wins $17 or loses $1. So the tickets in the box should be labeled ☐ and ☐.

• On each spin of the wheel, the gambler has _______ chances to win, and _______ chances to lose. So there should be _______ ☐'s and _______ ☐'s.

• Since there are 5 plays, there should be _______ draws.

• To illustrate the analogy:

Spin:	5	11	00	36	2		
Win/Loss:	___	___	___	___	___	Net gain:	___
Draws:	☐	☐	☐	☐	☐	Sum:	___

• In general, for gambling problems in which the same bet is made several times, a box model can be set up as follows:

- The tickets in the box should show the amounts that can be won (+) or lost (−) on each play.
- The chance of drawing any particular value from the box should equal the chance of winning that amount on a single play.
- The number of draws should equal the number of plays.
- Then the *net gain* is like the sum of the draws from the box.

• *Chuck-a-luck.* In one version of chuck-a-luck, 3 dice are rolled out of a cage. You can bet that all 3 show six. The house pays 36 to 1, and the bettor has ________ chance in ________ to win. Suppose you make this bet 10 times, staking $1 each time. Your net gain is like the sum of ________ draws made at random with replacement from the box

SUMMARY

• There is *chance error* in the number of heads:

number of heads = half the number of tosses + chance error

The error is likely to be large in absolute terms, but small relative to the number of tosses. This is the *law of averages*.

• The law of averages can be restated in percentage terms. With a large number of tosses, the percentage of heads is likely to be close to 50%, although it is not likely to be exactly equal to 50%.

• The law of averages does not work by compensation. When tossing a coin, for example, a run of heads is just as likely to be followed by a head as by a tail. A coin has no memory and no conscience.

• A complicated chance process for generating a number can often be modeled by drawing from a box. The sum of the draws is the key ingredient.

• The three basic questions to ask when making a box model are:

- What numbers go into the box?
- How many of each kind?
- How many draws?

- For gambling problems in which the same bet is made several times, a box model can be set up as follows:
    - The tickets in the box should show the amounts that can be won (+) or lost (−) on each play.
    - The chance of drawing any particular value from the box should equal the chance of winning that amount on a single play.
    - The number of draws should equal the number of plays.

Then the *net gain* is like the sum of the draws from the box.