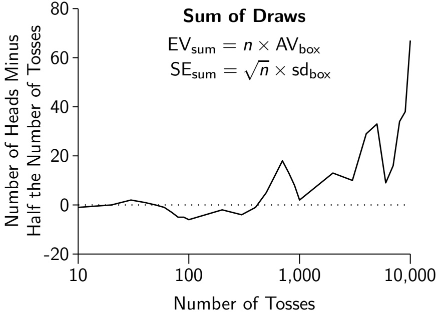
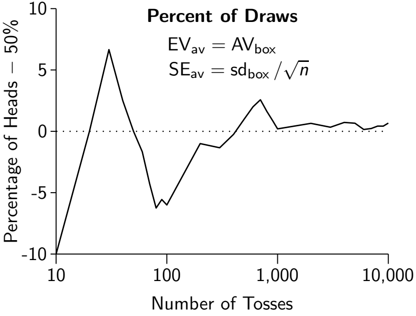
Math 207: Statistics

Chapter 17: Expected Value and Standard Error





Dr. Ralph Wojtowicz

Mathematics Department

Shenandoah University

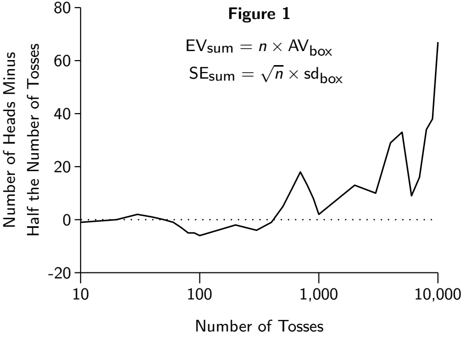
1. EV and SE
   - Law of Averages
   - EV and SE for Die Rolls
   - EV and SE for Coin Flips
   - EV and SE for Samples from a Box

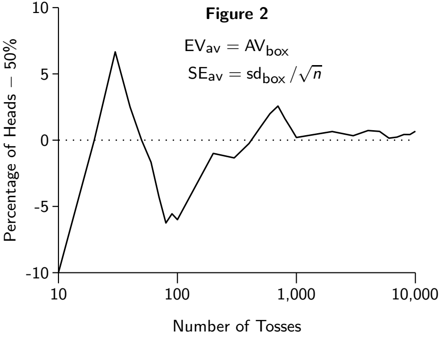
2. Using the Normal Curve
   - Using the Normal Curve
   - Examples I
   - Examples II

3. SD Shortcut
   - SD Shortcut

Law of Averages

The Law of Averages says that as the size of a random sample increases, the difference between the expected percentage of an outcome and the observed percentage will get smaller (Figure 2). The difference between the number of occurences of an outcome and the expected number increases (Figure 1).





Examples: Counting Dots on Dice

$$\mathrm{EV}_{\mathrm{sum}} = n \cdot \mathrm{AV}_{\mathrm{box}} \qquad \mathrm{EV}_{\mathrm{av}} = \mathrm{AV}_{\mathrm{box}}$$

$$\mathrm{SE}_{\mathrm{sum}} = \sqrt{n} \cdot \mathrm{SD}_{\mathrm{box}} \qquad \mathrm{SE}_{\mathrm{av}} = \mathrm{SD}_{\mathrm{box}}/\sqrt{n}$$

- Box model for rolling dice and taking the sum or average $\boxed{\boxed{1}\,\boxed{2}\,\boxed{3}\,\boxed{4}\,\boxed{5}\,\boxed{6}}$.
  $\mathrm{AV}_{\mathrm{box}} = 3.5$, $\mathrm{SD}_{\mathrm{box}} = \sqrt{35/12} \approx 1.71$

Examples: Counting Dots on Dice

$$\mathrm{EV}_{\mathrm{sum}} = n \cdot \mathrm{AV}_{\mathrm{box}} \qquad \mathrm{EV}_{\mathrm{av}} = \mathrm{AV}_{\mathrm{box}}$$

$$\mathrm{SE}_{\mathrm{sum}} = \sqrt{n} \cdot \mathrm{SD}_{\mathrm{box}} \qquad \mathrm{SE}_{\mathrm{av}} = \mathrm{SD}_{\mathrm{box}}/\sqrt{n}$$

- Box model for rolling dice and taking the sum or average $\boxed{\boxed{1}\,\boxed{2}\,\boxed{3}\,\boxed{4}\,\boxed{5}\,\boxed{6}}$.
  $\mathrm{AV}_{\mathrm{box}} = 3.5$, $\mathrm{SD}_{\mathrm{box}} = \sqrt{35/12} \approx 1.71$
  - Roll 10 dice: $\mathrm{EV}_{\mathrm{sum}} = 10 \cdot 3.5 = 35$ $\mathrm{SE}_{\mathrm{sum}} = 1.71 \cdot \sqrt{10} = 5.4$
    Simulated values: 26, 33, 45, 28, 37, 37, 36, 28, 35

Examples: Counting Dots on Dice

$$\text{EV}_{\text{sum}} = n \cdot \text{AV}_{\text{box}} \qquad \text{EV}_{\text{av}} = \text{AV}_{\text{box}}$$
$$\text{SE}_{\text{sum}} = \sqrt{n} \cdot \text{SD}_{\text{box}} \qquad \text{SE}_{\text{av}} = \text{SD}_{\text{box}}/\sqrt{n}$$

- Box model for rolling dice and taking the sum or average $\boxed{\boxed{1}\,\boxed{2}\,\boxed{3}\,\boxed{4}\,\boxed{5}\,\boxed{6}}$.
  $\text{AV}_{\text{box}} = 3.5$, $\text{SD}_{\text{box}} = \sqrt{35/12} \approx 1.71$
  - Roll 10 dice: $\text{EV}_{\text{sum}} = 10 \cdot 3.5 = 35$ $\text{SE}_{\text{sum}} = 1.71 \cdot \sqrt{10} = 5.4$
    Simulated values: 26, 33, 45, 28, 37, 37, 36, 28, 35
  - Roll 100 dice $\text{EV}_{\text{sum}} = 100 \cdot 3.5 = 350$ $\text{SE}_{\text{sum}} = 1.71 \cdot \sqrt{100} = 17.5$
    Simulated values: 326, 348, 374, 348, 361, 349, 361, 321

Examples: Counting Dots on Dice

$$\mathrm{EV}_{\mathrm{sum}} = n \cdot \mathrm{AV}_{\mathrm{box}} \qquad \mathrm{EV}_{\mathrm{av}} = \mathrm{AV}_{\mathrm{box}}$$

$$\mathrm{SE}_{\mathrm{sum}} = \sqrt{n} \cdot \mathrm{SD}_{\mathrm{box}} \qquad \mathrm{SE}_{\mathrm{av}} = \mathrm{SD}_{\mathrm{box}}/\sqrt{n}$$

- Box model for rolling dice and taking the sum or average $\boxed{\boxed{1}\,\boxed{2}\,\boxed{3}\,\boxed{4}\,\boxed{5}\,\boxed{6}}$.
  $\mathrm{AV}_{\mathrm{box}} = 3.5$, $\mathrm{SD}_{\mathrm{box}} = \sqrt{35/12} \approx 1.71$
  - Roll 10 dice: $\mathrm{EV}_{\mathrm{sum}} = 10 \cdot 3.5 = 35$ $\mathrm{SE}_{\mathrm{sum}} = 1.71 \cdot \sqrt{10} = 5.4$
    Simulated values: 26, 33, 45, 28, 37, 37, 36, 28, 35
  - Roll 100 dice $\mathrm{EV}_{\mathrm{sum}} = 100 \cdot 3.5 = 350$ $\mathrm{SE}_{\mathrm{sum}} = 1.71 \cdot \sqrt{100} = 17.5$
    Simulated values: 326, 348, 374, 348, 361, 349, 361, 321
  - Roll 10 dice: $\mathrm{EV}_{\mathrm{av}} = 3.5$ $\mathrm{SE}_{\mathrm{av}} = 1.71/\sqrt{10} = 0.54$
    Simulated values: 4.1, 3.7, 3.6, 3.6, 3.9, 3.2, 3.1, 3.8

Examples: Counting Dots on Dice

$$\text{EV}_{\text{sum}} = n \cdot \text{AV}_{\text{box}} \qquad \text{EV}_{\text{av}} = \text{AV}_{\text{box}}$$

$$\text{SE}_{\text{sum}} = \sqrt{n} \cdot \text{SD}_{\text{box}} \qquad \text{SE}_{\text{av}} = \text{SD}_{\text{box}}/\sqrt{n}$$

- Box model for rolling dice and taking the sum or average $\boxed{\boxed{1}\ \boxed{2}\ \boxed{3}\ \boxed{4}\ \boxed{5}\ \boxed{6}}$.
  $\text{AV}_{\text{box}} = 3.5$, $\text{SD}_{\text{box}} = \sqrt{35/12} \approx 1.71$
  - Roll 10 dice: $\text{EV}_{\text{sum}} = 10 \cdot 3.5 = 35$ $\text{SE}_{\text{sum}} = 1.71 \cdot \sqrt{10} = 5.4$
    Simulated values: 26, 33, 45, 28, 37, 37, 36, 28, 35
  - Roll 100 dice $\text{EV}_{\text{sum}} = 100 \cdot 3.5 = 350$ $\text{SE}_{\text{sum}} = 1.71 \cdot \sqrt{100} = 17.5$
    Simulated values: 326, 348, 374, 348, 361, 349, 361, 321
  - Roll 10 dice: $\text{EV}_{\text{av}} = 3.5$ $\text{SE}_{\text{av}} = 1.71/\sqrt{10} = 0.54$
    Simulated values: 4.1, 3.7, 3.6, 3.6, 3.9, 3.2, 3.1, 3.8
  - Roll 100 dice: $\text{EV}_{\text{av}} = 3.5$ $\text{SE}_{\text{av}} = 1.71/\sqrt{100} = 0.171$
    Simulated values: 3.53, 3.34, 3.53, 3.24, 3.55, 3.58, 3.69, 3.31

Examples: Counting Coin Flips

$$\text{EV}_{\text{sum}} = n \cdot \text{AV}_{\text{box}} \qquad \text{EV}_{\text{av}} = \text{AV}_{\text{box}}$$

$$\text{SE}_{\text{sum}} = \sqrt{n} \cdot \text{SD}_{\text{box}} \qquad \text{SE}_{\text{av}} = \text{SD}_{\text{box}}/\sqrt{n}$$

- Box model for flipping coins and taking the sum or average $\boxed{\boxed{0}\ \boxed{1}}$.
  $\text{AV}_{\text{box}} = 0.5$, $\text{SD}_{\text{box}} = 0.5$.

Examples: Counting Coin Flips

$$\text{EV}_{\text{sum}} = n \cdot \text{AV}_{\text{box}} \qquad \text{EV}_{\text{av}} = \text{AV}_{\text{box}}$$

$$\text{SE}_{\text{sum}} = \sqrt{n} \cdot \text{SD}_{\text{box}} \qquad \text{SE}_{\text{av}} = \text{SD}_{\text{box}}/\sqrt{n}$$

- Box model for flipping coins and taking the sum or average $\boxed{\boxed{0}\ \boxed{1}}$.
  $\text{AV}_{\text{box}} = 0.5$, $\text{SD}_{\text{box}} = 0.5$.
  - Flip 10 coins: $\text{EV}_{\text{sum}} = 10 \cdot 0.5 = 5$ $\text{SE}_{\text{sum}} = 0.5 \cdot \sqrt{10} = 1.58$
    Simulated values: 5, 4, 6, 5, 2, 6, 5, 5

Examples: Counting Coin Flips

$$EV_{sum} = n \cdot AV_{box} \qquad EV_{av} = AV_{box}$$
$$SE_{sum} = \sqrt{n} \cdot SD_{box} \qquad SE_{av} = SD_{box}/\sqrt{n}$$

- Box model for flipping coins and taking the sum or average $\boxed{\boxed{0}\ \boxed{1}}$.
  $AV_{box} = 0.5$, $SD_{box} = 0.5$.
  - Flip 10 coins: $EV_{sum} = 10 \cdot 0.5 = 5$ $SE_{sum} = 0.5 \cdot \sqrt{10} = 1.58$
    Simulated values: 5, 4, 6, 5, 2, 6, 5, 5
  - Flip 100 coins $EV_{sum} = 100 \cdot 0.5 = 50$ $SE_{sum} = 0.5 \cdot \sqrt{100} = 5$
    Simulated values: 47, 48, 48, 42, 53, 54, 52, 45

Examples: Counting Coin Flips

$$\text{EV}_{\text{sum}} = n \cdot \text{AV}_{\text{box}} \qquad \text{EV}_{\text{av}} = \text{AV}_{\text{box}}$$

$$\text{SE}_{\text{sum}} = \sqrt{n} \cdot \text{SD}_{\text{box}} \qquad \text{SE}_{\text{av}} = \text{SD}_{\text{box}}/\sqrt{n}$$

- Box model for flipping coins and taking the sum or average $\boxed{\boxed{0}\ \boxed{1}}$. $\text{AV}_{\text{box}} = 0.5$, $\text{SD}_{\text{box}} = 0.5$.
  - Flip 10 coins: $\text{EV}_{\text{sum}} = 10 \cdot 0.5 = 5$ $\text{SE}_{\text{sum}} = 0.5 \cdot \sqrt{10} = 1.58$
    Simulated values: 5, 4, 6, 5, 2, 6, 5, 5
  - Flip 100 coins $\text{EV}_{\text{sum}} = 100 \cdot 0.5 = 50$ $\text{SE}_{\text{sum}} = 0.5 \cdot \sqrt{100} = 5$
    Simulated values: 47, 48, 48, 42, 53, 54, 52, 45
  - Flip 10 coins: $\text{EV}_{\text{av}} = 0.5$ $\text{SE}_{\text{av}} = 0.5/\sqrt{10} = 0.16$
    Simulated values: 0.2, 0.4, 0.5, 0.1, 0.6, 0.6, 0.4, 0.6

Examples: Counting Coin Flips

$$\mathrm{EV}_{\mathrm{sum}} = n \cdot \mathrm{AV}_{\mathrm{box}} \qquad \mathrm{EV}_{\mathrm{av}} = \mathrm{AV}_{\mathrm{box}}$$

$$\mathrm{SE}_{\mathrm{sum}} = \sqrt{n} \cdot \mathrm{SD}_{\mathrm{box}} \qquad \mathrm{SE}_{\mathrm{av}} = \mathrm{SD}_{\mathrm{box}}/\sqrt{n}$$

- Box model for flipping coins and taking the sum or average $\boxed{\boxed{0}\ \boxed{1}}$.
  $\mathrm{AV}_{\mathrm{box}} = 0.5$, $\mathrm{SD}_{\mathrm{box}} = 0.5$.
  - Flip 10 coins: $\mathrm{EV}_{\mathrm{sum}} = 10 \cdot 0.5 = 5$ $\mathrm{SE}_{\mathrm{sum}} = 0.5 \cdot \sqrt{10} = 1.58$
    Simulated values: 5, 4, 6, 5, 2, 6, 5, 5
  - Flip 100 coins $\mathrm{EV}_{\mathrm{sum}} = 100 \cdot 0.5 = 50$ $\mathrm{SE}_{\mathrm{sum}} = 0.5 \cdot \sqrt{100} = 5$
    Simulated values: 47, 48, 48, 42, 53, 54, 52, 45
  - Flip 10 coins: $\mathrm{EV}_{\mathrm{av}} = 0.5$ $\mathrm{SE}_{\mathrm{av}} = 0.5/\sqrt{10} = 0.16$
    Simulated values: 0.2, 0.4, 0.5, 0.1, 0.6, 0.6, 0.4, 0.6
  - Flip 100 coins: $\mathrm{EV}_{\mathrm{av}} = 0.5$ $\mathrm{SE}_{\mathrm{av}} = 0.5/\sqrt{100} = 0.05$
    Simulated values: 0.50, 0.54, 0.54, 0.47, 0.56, 0.38, 0.43, 0.53

Example: Sampling from a Box

- One hundred draws are to be made at random with replacement from the box

$$\boxed{\boxed{1}\ \boxed{2}\ \boxed{2}\ \boxed{5}\ \boxed{9}\ \boxed{10}}$$

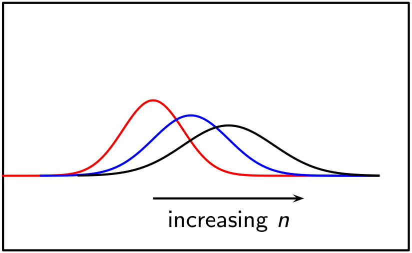
- $\text{AV}_{\text{box}} = \frac{1+2+2+5+9+10}{6} \approx 4.83$
- $\text{SD}_{\text{box}} = \sqrt{\frac{(1-4.83)^2 + (2-4.83)^2 + \cdots + (10-4.83)^2}{6}} \approx 3.53$
- $\text{EV}_{\text{sum}} = n \cdot \text{AV}_{\text{box}} \approx 100 \cdot 4.83 = 483.$
- $\text{SE}_{\text{sum}} = \sqrt{n} \cdot \text{SD}_{\text{box}} \approx 10 \cdot 3.53 = 35.3.$
- The sum of the draws will be around 483 give or take 35.3 or so.
- Simulated values: 442, 512, 530, 515, 464, 545, 481, 482
- $\text{EV}_{\text{av}} = \text{AV}_{\text{box}} \approx 4.83.$
- $\text{SE}_{\text{av}} = \text{SD}_{\text{box}}/\sqrt{n} \approx 3.53/10 = 0.353.$
- The average of the draws will be around 4.83 give or take 0.353 or so.
- Simulated values: 4.93, 5.01, 4.62,4.44, 5.08, 4.98, 4.37, 5.21

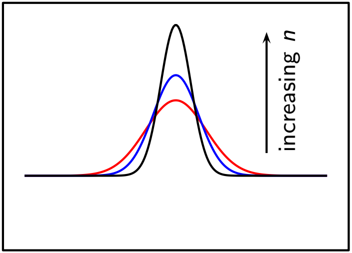
Using the Normal Curve

Central Limit Theorem: When drawing at random with replacement from a box, the probability histogram for the sum (and the average) will follow the normal curve, even if the contents of the box do not. The histogram must be put into standard units, and the number of draws must be reasonably large.

$$\mathrm{EV_{sum}} = n \cdot \mathrm{AV_{box}} \qquad \mathrm{EV_{av}} = \mathrm{AV_{box}}$$

$$\mathrm{SE_{sum}} = \sqrt{n} \cdot \mathrm{SD_{box}} \qquad \mathrm{SE_{av}} = \frac{\mathrm{SD_{box}}}{\sqrt{n}}$$





Examples (I)

- A coin will be flipped 100 times. You are about 68% confident that the number of tails will fall in what range?
  The box is $\boxed{\boxed{0}\ \boxed{1}}$.
  $AV_{box} = 0.5$ and $SD_{box} = 0.5$.
  $EV_{sum} = 0.5 \cdot 100 = 50$ and $SE_{sum} = 0.5 \cdot \sqrt{100} = 5$
  Answer: 50 ± 5
- A coin will be flipped 10,000 times. You are about 95% confident that the number of tails will fall in what range?
  $EV_{sum} = 0.5 \cdot 10{,}000 = 5000$ and $SE_{sum} = 0.5 \cdot \sqrt{10000} = 50$
  Answer: $5000 \pm 2 \cdot 50 = 5000 \pm 100$
- A coin will be flipped 10,000 times. You are about 95% confident that the percent of tails will fall in what range?
  $EV_{av} = 0.5 = 50\%$ and $SE_{av} = 0.5/\sqrt{10000} = 0.005 = 0.5\%$
  Answer: $50\% \pm 2 \cdot 0.5\% = 50\% \pm 1\%$

Examples (II)

Examples (II)

- A die will be tossed 120 times. What is the chance that the number of 4s will be between 15 and 25?

  The box model is $\boxed{\boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{1}\ \boxed{0}\ \boxed{0}}$

  $\text{AV}_{\text{box}} = \frac{1}{6}$ and $\text{SD}_{\text{box}} = \frac{\sqrt{5}}{6} \approx 0.373.$

  $\text{EV}_{\text{sum}} = 120 \cdot \frac{1}{6} = 20$ and $\text{SE}_{\text{sum}} = \sqrt{120} \cdot \frac{\sqrt{5}}{6} \approx 4.08$

  $z_1 = \frac{15-20}{4.08} = -1.23$ and $z_2 = \frac{25-20}{4.08} = 1.23$

  Chance $= \text{pnorm}(1.23) - \text{pnorm}(-1.23) = 0.781 = 78.1\%$

Examples (II)

- A die will be tossed 120 times. What is the chance that the number of 4s will be between 15 and 25?

  The box model is $\boxed{\boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{1}\ \boxed{0}\ \boxed{0}}$

  $\text{AV}_{\text{box}} = \frac{1}{6}$ and $\text{SD}_{\text{box}} = \frac{\sqrt{5}}{6} \approx 0.373$.

  $\text{EV}_{\text{sum}} = 120 \cdot \frac{1}{6} = 20$ and $\text{SE}_{\text{sum}} = \sqrt{120} \cdot \frac{\sqrt{5}}{6} \approx 4.08$

  $z_1 = \frac{15-20}{4.08} = -1.23$ and $z_2 = \frac{25-20}{4.08} = 1.23$

  Chance $=$ pnorm$(1.23) -$ pnorm$(-1.23) = 0.781 = 78.1\%$

- A die will be tossed 12,000 times. What is the chance that the percentage of 4s will be between 16% and 17%?

  $\text{EV}_{\text{av}} = \frac{1}{6} = 0.166667 = 16.7\%$ and $\text{SE}_{\text{av}} = \frac{\sqrt{5}}{6}/\sqrt{12000} \approx 0.0034 = 0.34\%$

  $z_1 = \frac{16-16.7}{0.34} = -1.96$ and $z_2 = \frac{17-16.7}{0.34} = 0.98$

  Chance $=$ pnorm$(0.98) -$ pnorm$(-1.96) = 0.811 = 81.1\%$

SD Shortcut

- If there are only two different numbers on the tickets in a box then we can use a shortcut formula to compute the SD of the box.

$$\mathrm{SD}_{\text{box}} = \left(\begin{array}{c}\text{big}\\ \text{number}\end{array} - \begin{array}{c}\text{small}\\ \text{number}\end{array}\right)\sqrt{\left(\begin{array}{c}\text{fraction of}\\ \text{tickets with}\\ \text{big number}\end{array}\right)\cdot\left(\begin{array}{c}\text{fraction of}\\ \text{tickets with}\\ \text{small number}\end{array}\right)}$$

- Examples:

	Box	SD
Flipping coins	0 1	$(1-0)\sqrt{\frac{1}{2}\cdot\frac{1}{2}} = \frac{1}{2} = 0.5$
Counting 4s on die rolls	0 0 0 1 0 0	$(1-0)\sqrt{\frac{1}{6}\cdot\frac{5}{6}} = \frac{\sqrt{5}}{6} \approx 0.373$
Sample from a box	-1 -1 1 1 1	$(1-(-1))\sqrt{\frac{3}{5}\cdot\frac{2}{5}} = \frac{2\sqrt{6}}{5} \approx 0.980$
Sample from a box	1 3 3 5	We can't use the shortcut.