this time: probability models for sums & means

read: DD (?) ch. 11

AMS 7
5 Jul 2016

1

next time: statistical models for means; interval estimation

LN pp. 137-160

today: LN pp. 119 → 127

take home midterm due next Thu in disc. sec.

homework 1 grades posted on ecommons

mean 88/100
median 87
SD 11

score	%
100+	24
90-99	18
80-89	41
70-79	12
60-69	6

P(coming out ahead on any single play with strategy A) $= \frac{1}{38} \doteq 2.5\%$

P(— strategy B) $= \frac{2}{38} = \frac{1}{19} \doteq 5\%$

$$\mu = \frac{\overbrace{(\overset{\$}{-1}) + (\overset{\$}{-1}) + \dots + (\overset{\$}{-1})}^{37} + (\overset{\$}{+35})}{38}$$

$$= \frac{37(\overset{\$}{-1}) + 1(\overset{\$}{+35})}{38} = \frac{\overset{\$}{-2}}{38} = -\frac{\overset{\$}{1}}{19} \doteq \overset{\$}{-0.0526}$$

pop. — all possibilities on any one spin

single # — A (6)

sample — the observed spins

imaginary data set — all possible samples

your net gain

#	$
1	−1
2	−1
3	⋮
4	⋮
5	−1
6	$+35
7	−1
⋮	⋮
36	−1
0	−1
00	−1

$N = 38$

"new"

mean $\mu \doteq -\$0.05$

SD $\sigma = \$5.76$

"sigma"

at random with repl = IID

your net gain

$ −1
$+35
−1
⋮
−1

$n = 1{,}000$

sum $ = ?

ex. −$61

IID

$n = 1{,}000$

sum $ = ?

(ex $+11)

$ −61
+11
⋮
⋮

$M \to \infty$

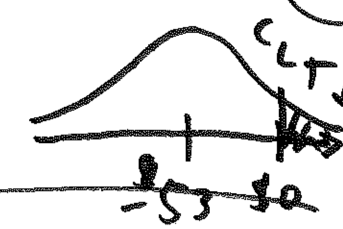
long-run mean — EV of $ = −$52.60

long-run SD — SE of $ ≐ $182

long-run hist. — SE $182

39%

CLT



(6.52)

$$+0.29 = \frac{53}{182} = \frac{(\$0) - (-\$53)}{\$182}$$

P(coming out ahead (strategy A)) = P($ > $0) ≐ 39%

every time wheel spins with strategy A,
I expect to win or lose around $\underline{\mu = \$-0.05}$,
give or take around $\underline{\sigma \doteq \$5.76}$

$$\sigma = \sqrt{\frac{\left[\$-1 - (\$-0.05)\right]^2 + \dots + \left[\$-1 - (\$-0.05)\right]^2 + \left[(\$+35) - (\$-0.05)\right]^2}{38}}$$

---

math fact:

with any population having only 2 possible values, the pop. SD is

$$\sigma = \left[\begin{pmatrix}\text{larger}\\\text{value}\end{pmatrix} - \begin{pmatrix}\text{smaller}\\\text{value}\end{pmatrix}\right]\sqrt{p \cdot (1-p)}$$

p = proportion of larger values

here, smaller value = -1

larger value = $+35$

$$p = \frac{1}{38} \qquad 1-p = \frac{37}{38}$$

so $$\sigma = \underbrace{\left[(+\$35) - (-\$1)\right]}_{(+\$36)} \sqrt{\frac{1}{38} \cdot \frac{37}{38}} \doteq 85.76$$

~~real world~~ →

your net gain after $n =$ 1,000 \$1 bets on a single # is like the sum of $n = 1{,}000$ IID draws from the single-# population ← the model long-run mean of sums S is imag. dataset = expected value of the sum S = EV of S

$$= \boxed{E_{IID}(S) = n\mu} = \left(\begin{matrix}\#\\ \text{draws}\end{matrix}\right)\left(\begin{matrix}\text{pop.}\\ \text{mean}\end{matrix}\right)$$

math fact

$$= (1{,}000)(-\$0.0526) \doteq -\$52.60$$

---

at the end of $n = 1{,}000$ \$1 bets on a single #, I expect to be ahead (behind) by EV = −\$52.60, give or take about SE ≐ \$182 long-run SD of sum $ in imag. data set = <u>standard error</u> of the sum $ = SE of $

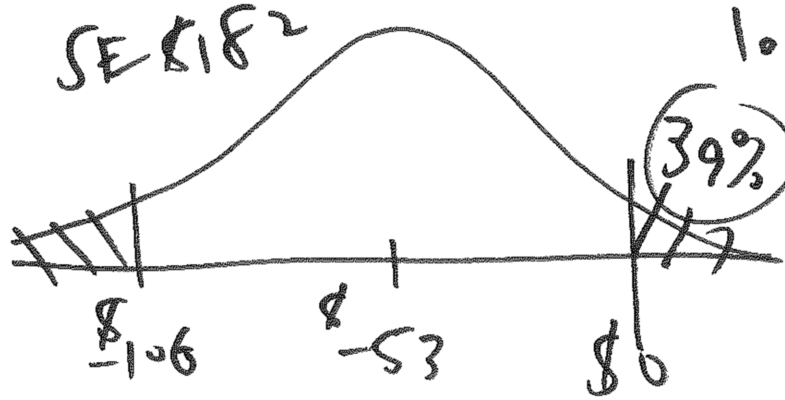
$$= \boxed{SE_{IID}(\$) = \sigma \cdot \sqrt{n}} = \begin{pmatrix}\text{pop}\\ \text{SD}\end{pmatrix} \cdot \sqrt{\#\text{ draws}}$$

N	X
μ	X
σ ↑	$SE(\$)$ ↑
n ↑	$SE(\$)$ ↑

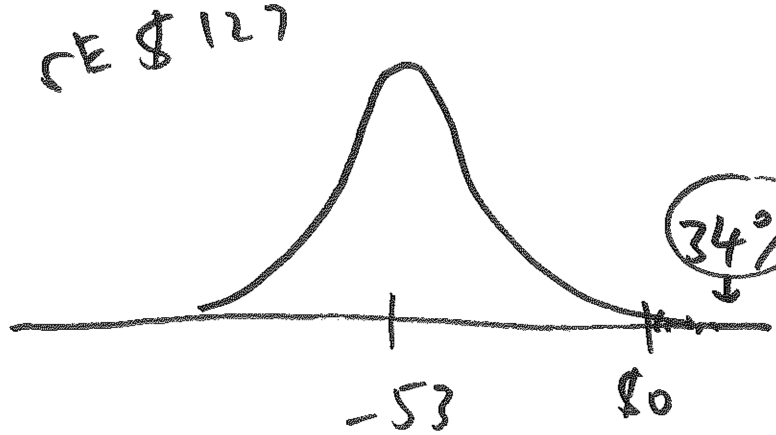
$$SE_{IID}(\$) = \frac{\sigma \cdot \sqrt{n}}{1} = \sigma\sqrt{n}$$

here $SE = (\$5.76)\sqrt{1000}$

$\doteq \$182$



long run hist of $ (single #)



long run hist of $ (split)

$$\frac{(\$0) - (\$53)}{\$127} \doteq +.42$$