10/25/18

Wave Moretto

Lecture # 9: Probability Models for sums

P(1 or more T-S babies in family of 5, both parents carriers)

$= 1 - P(0 \text{ T-S babies})$

$= 1 - P\left(\begin{matrix}\text{not T-S} \\ \text{on first}\end{matrix} \text{ and } \begin{matrix}\text{not T-S} \\ \text{on 2nd}\end{matrix} \text{ and } \ldots \begin{matrix}\text{not T-S} \\ \text{on 5th}\end{matrix}\right)$

IID

$= 1 - P\left(\begin{matrix}\text{not} \\ \text{T-S 1st}\end{matrix}\right) \cdot P\left(\begin{matrix}\text{not} \\ \text{T-S 2nd}\end{matrix}\right) \cdot \ldots P\left(\begin{matrix}\text{not} \\ \text{T-S 5th}\end{matrix}\right)$

$= 1 - \frac{3}{4}\left(\frac{3}{4}\right) \ldots \left(\frac{3}{4}\right)$

$= 1 - \left(\frac{3}{4}\right)^5 = \boxed{76\%}$

*useful for problem on fakenume midterm

T-sac babies
0
1
2
3
4
5

1 or more

if ELM applies here

$P(1 \text{ or more}) = \frac{5}{6}$

but ELM doesn't apply

- see ven. diagram p. L-104

New Unit: Probability Models for sums

R-52

L-119

- **Roulette**

$P(\text{coming out ahead, 1 \$1 bet on A}) = \frac{1}{38}$

ELM?

$= 2.6\%$

$P(\text{coming out ahead, 1 \$1 bet on B}) = \frac{2}{38}$

$= 5.2\%$

Pop. — all possible spins — your net gain

N = 38

$\begin{bmatrix} -\$1 \\ -\$1 \\ -\$1 \\ \vdots \\ +\$35 \\ \vdots \end{bmatrix}$ 2, 3, 4 … 36, 0, 00

wheel fair

IID →

$M = -\$0.05$

$SD = \$5.76$

Sample — the observed spins — your net gain

$\begin{bmatrix} -\$1 \\ -\$1 \\ +\$35 \\ -\$1 \\ \vdots \end{bmatrix}$ $n = 1{,}000$

sum \$ = ?

ex: -64

imaginary data set / repeated sampling data set

$[\$]^{n=1000}$

1st = ?

ex: +8

$\begin{bmatrix} \$ \\ -64 \\ +8 \\ \vdots \end{bmatrix}$

$M \to \infty$

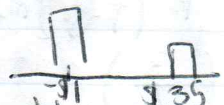
P(coming out ahead Ⓐ) = P($S > \$0$) = ?

(*) Pop. mean μ "mew" = $\frac{(-1)+(-1)+\ldots(-1)+35}{38} = \frac{37(-1)+1(+35)}{38}$

$= \frac{-\$2}{38} = -\$0.05$

pop. SD σ sigma $= \$5.76$

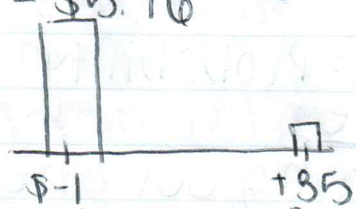
pop histogram:



- With every \$1 bet on a single number, it expect to loose μ \$0.05, give or take σ \$5.76
- Math fact: if pop. has only 2 values in it, σ = pop SD = (large value − smaller value) $\cdot$ $\sqrt{(\text{prop larger values})(\text{prop smaller values})}$
- here $\sigma = [(+35)-(-1)]\sqrt{\frac{1}{38}\cdot\frac{37}{38}}$

$= \$5.76$

- Histogram:



- Your net gain after 1000 \$1 bets on a single # (real world) is exactly like the sum S of n=1000 IID draws from this (*) population (math + computations)
  - probability model ↑ approx simulation ↑
  - essence of applied math (1950)

range of possible values for S

-\$1000 π -\$53 ± 2(182) \$35,000

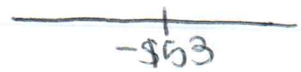
$+\$35\left(\frac{1}{38}\right)$ $\frac{1000}{38} = 26$ $26(+\$35) + 974(-1)$

$= -\$64$

$-\$1\left(\frac{37}{38}\right)$ $28(+35)+972(-1) = +8$

$(\mu \to \infty)$

- Long run mean: $-\$53$
- Long run SD: $\$182$
- Long run hist.



- ~~Long run mean~~ = expected value of $ = EV of $

$= E_{IID}(\$) = n \cdot \mu = (\#\text{ draws}) \cdot (\text{pop. mean})$

$= (1000)(-0.0526) = -\$52.60$

- ~~Long run SD of $~~ = standard error of $ =

$SE_{IID}(\$) = \text{SE of }\$$ ← give or take for $

"noise level in predicting $"

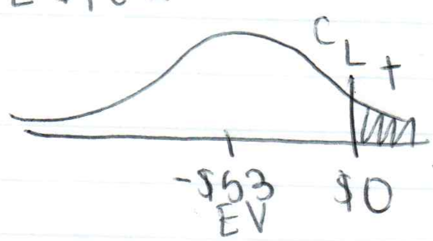
N	X
μ	X
σ	as $\sigma\uparrow$ $SE(\$)\uparrow$
n	as $n\uparrow$ $SE(\$)\uparrow$

$\boxed{SE_{IID}(\$) = \sigma n}$ / 1 not right... $\Rightarrow SE_{IID}(\$) = \sigma\sqrt{n}$

- after 1000 $1 bets on a single #, I expect to have gained ($-53)
- $SE_{IID} = \sigma\sqrt{n} = (\text{pop SD})\sqrt{\#\text{ draws}}$

$= (\$5.76)\sqrt{1000} = \boxed{\$182}$

SE $182



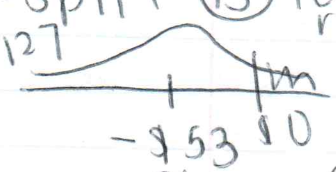
** long run hist. of $

- ~~central limit theorem~~
- (L-124) $n \to \infty$

39% SE $127

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

- Split (13) long run hist. of $



$\frac{0-(-53)}{127} = 0.42$ 34%