Julia Larson
10/20
AMS 7
Notes
1

10/20 Probability models for sums and means cont.

Your net gain after a 1000 \$1 bets on a single number $\boxed{\text{is like}}$ the sum of $n=1000$ IID draws of that population.

Roulette cont.

What is the standard error of S = SE = $SE_{IID}(S)$ =
(Long run SD)

N = not in this formula

μ = not in this formula

σ = As $\sigma\uparrow$ SE $\uparrow$

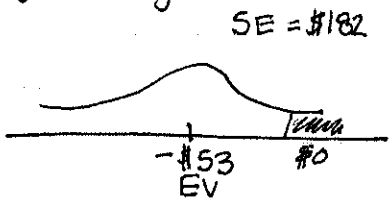
n = As $n\uparrow$ $SE(S)\uparrow$

Math Fact → $SE_{IID}(S) = \frac{\sigma\sqrt{n}}{1} = \sigma\sqrt{n} = (\text{Pop SD})\sqrt{\#\text{ of draws}}$

Here $SE(S) = \$5.76\sqrt{1000} = \$182$

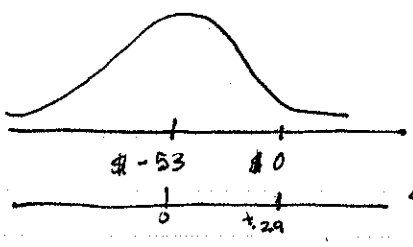
After a 1000 \$1 plays of single # I expect to be behind by about \$53 (EV) give or take about \$182 (SE)

Long range histogram of the sum (S)



Central limit theorem (CLT): as long as n is large the long run hist. of sum (or mean) of (n) IID draws from a pop. will look a lot like the normal curve pretty much no matter what the population histogram looks like, the closer the pop. histogram is to the normal curve to begin with the smaller (n) needs to be to get a good normal approximation if the pop. histogram is normal to begin with, then the long-run hist. of S (or mean) in imaginary data set is normal for all $(n \geq 1)$

long run histogram of S (single #)



$\frac{0-(\$-53)}{\$182} = z$ = .3351 = 34%

Julia Larsen
10/20
AMS 7
Notes
2

Same process for split

Pop: ditto

Sample: ditto

ditto

$N = 38$: $+\$17$, $+\$17$, $-\$1$, $-\$1$, $-\$1$, $-\$1$ (2, 36)

IID → ditto [ ] $n = 1000$

Sum = ? S (ex. $-\$54$)

IID → [⋮]

Sum S = ?

⋮

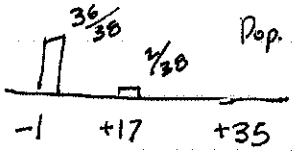
ditto

[$\$-54$, $\$+8$, ⋮] $m = \infty$

mean $\mu = \$-.05$

SD $= \sigma = \$4.02$

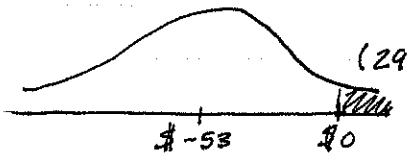
Pop. Histogram



long run mean $= -\$53$

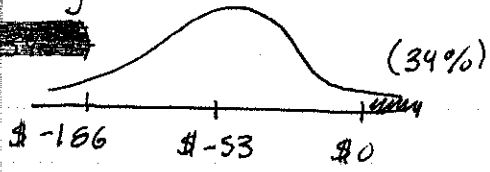
long run SD $= \$127$

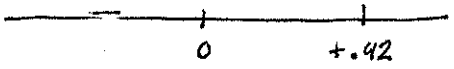
Split #



$\frac{\$0 - -53}{182} = .2912 = 29\%$

Single #





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

So there is a higher chance of coming out ahead with a single bet. Bold people bet singles, cautiouse ones bet splits.

In general bolder play works best in games with higher SD's.

Measuring Butter

$\begin{bmatrix} 16\,oz \\ 16\,oz \\ 16\,oz \\ 16\,oz \end{bmatrix}$ or $\begin{bmatrix} 16.0\,oz \\ 16.0\,oz \\ 16.0\,oz \\ 16.0\,oz \end{bmatrix}$ $\begin{bmatrix} 15.97\,oz \\ 16.01\,oz \\ \vdots \\ 15.99\,oz \end{bmatrix}$

deterministic

Stochostic (random)

There is always a point where things switch from deterministic to stochastic this is called measurement error.

Julia Larson
10/20
AMS 7
Notes
3

Basic Measurment error model

observable		Not observable		Random Error (IID) Mean = 0
Observation 1	=	(True Value)	+ (bias) +	"error1"
Observation 2	=	(True Value)	+ (bias) +	"error2"
⋮		⋮	⋮	⋮
Observation n	=	(True Value)	+ (bias) +	"error n"

Good data-gathering activity: unbiased (bias = 0)

Suppose truth = 3.8 (not hypo Kolemic) * measuring process unbiased

$3.9 = 3.8 + 0 + (+0.1)$

$3.5 = 3.8 + 0 + (-0.3)$

⋮

$4.0 = 3.8 + 0 + (+0.2)$

$$\text{mean of n obs.} = 3.8 + 0 + \frac{(+0.1) + (-0.3) + \dots + (+0.2)}{n}$$

mean of n random error with mean 0

There is cancellation of ⊕ & ⊖ numbers mean if n IID random error (each with mean 0) will likely be a lot smaller than any of the individual errors themselves

Unbiased true value 3.8

Pop	Sample	Imaginary data
Conceptual: all possible measurements	observed measurements	

if N = ∞ SRS is the same as IID

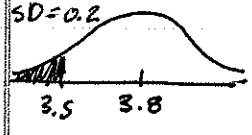
Pottasium — Like SRS → Potassium

N = ∞ $\begin{bmatrix} \\ \\ \\ \end{bmatrix}$ → $\begin{bmatrix} Y_1 \\ \vdots \\ Y_4 \end{bmatrix}$ n = 4 $\begin{bmatrix} \\ \\ \end{bmatrix}$

mean $\bar{Y} = ?$

mean $\mu = 3.8$

SD $\sigma = 0.2$



Pop. Hist. = hist. of 1 reading at a time

P(misdiagnosis with n=1) = 7% Bad.