

this time: sampling, probability

14 ①  
Apr

Ans  
5

next time: probability

read:  
) Dec. 8

Fpp  
ch. 14

T: iso (blue pill)

C: placebo (blue pill)

} blinding subjects  
to their T/C  
status

good idea to blind investigators to  
T/C status for all subjects too

both blinded: double-blind

after controlling for age; it  
appears that taking pill causes an  
increase in b.p. of about 5 mmHg

III. B

Sample Surveys:

population

entire collection of subjects (people, things, ...) of interest to you

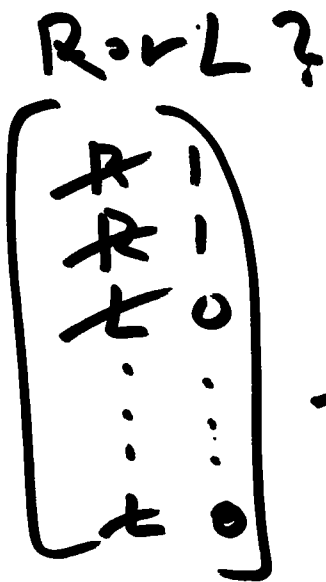
pop in '86:

\* everybody who was going to show up to vote in Nov '86

$N = \# \text{ elements/subjects in pop}$   
 $\approx \text{pop. size} = 100,000,000$

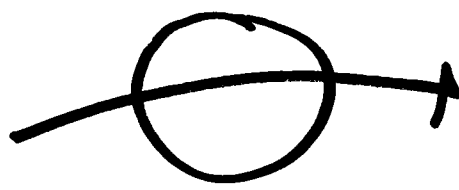
pop

$1=R$   
 $0=L$   
 $N=100,000,000$



sample  
the observed people

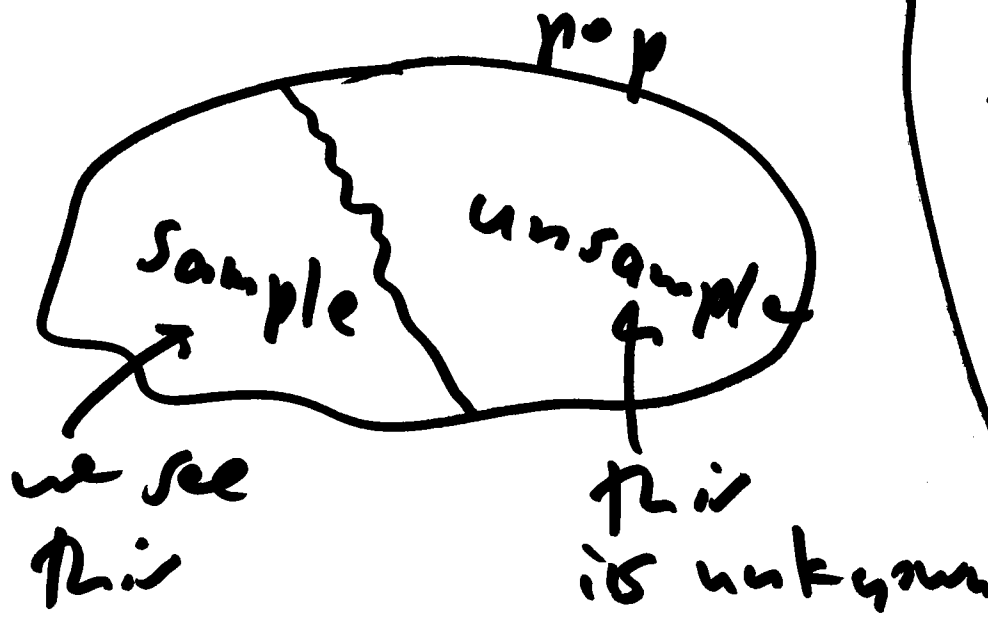
subset of  $pop$



L.D.  
 $n = 2,400,000$   
sample size

mean % for  $R = p = ?$   
in pop.

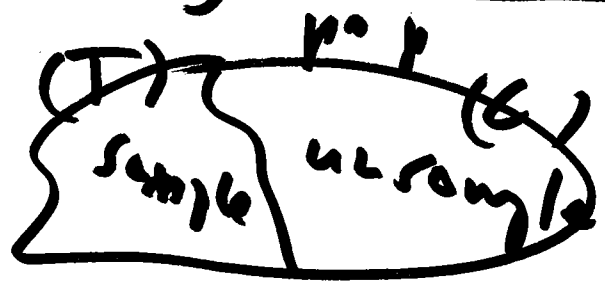
mean  $\bar{y} = \hat{p} = 43\%$



pop summary of interest: parameter (here parameter is  $p$ )

sample summary used to estimate (guess at)  $p$  :  $\hat{p}$  ( $p$ -hat) (estimate)

goal in sampling: make sample, unsample similar in relevant ways



a simple method to promote similarity of sample, unsample: choose sample at random

at random  
with  
replacement

independent  
identically  
distributed (IID)  
sampling

at random  
without  
replacement

simple random  
sampling (SRS)

$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$



$\begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix}$

IID:  
math  
easier

SRS: more informative than  
IID but math harder

if pop. size  $N$  is a lot

larger than sample size  $n$

$(N \gg n)$  : SRS  $\doteq$  IID

what  
people do

use this  
for math

