This time: sampling probability

next time: probability read DD ch. 8

T: iso (blue pill) blinding subjects to pair T/C status
C: placebo (blue pill) 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:
The sample of people selected for the opinion poll is chosen.

The opinion poll is conducted.

The mean % for K = p = ?

N = 0.01 \times 1000 = 100

R = 10 + 0

p = \frac{\text{number of people who voted for } K}{N}

\text{Sample of people selected for the opinion poll is chosen.}

\text{The opinion poll is conducted.}

\text{The mean } \% \text{ for } K = p = ?

\text{N} = 0.01 \times 1000 = 100

\text{R} = 10 + 0

p = \frac{\text{number of people who voted for } K}{N}
we see this

This is unknown

Sample

Unsample

Population summary of interest: parameter

(Here parameter is \( p \))

Sample summary used to estimate (guess at) \( \hat{p} \) (\( p - \text{hat} \)) (estimate)

Goal in sampling:

Make sample, unsample similar in all relevant ways

A simple method to promote similarity of sample, unsample:

Choose sample at random
<table>
<thead>
<tr>
<th>at random with replacement</th>
<th>independent identically distributed (IID) sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>at random without replacement</td>
<td>simple random sampling (SRS)</td>
</tr>
</tbody>
</table>

\[
\begin{bmatrix}
1 \\
2 \\
9
\end{bmatrix} \rightarrow \begin{bmatrix}
\end{bmatrix}
\]

IID: math easier

SRS: more informative than IID but math harder

If pop. size $N$ is a lot bigger than sample size $n$ ($N \gg n$): SRS $\approx$ IID

What people do for math
Fisher (randomness, exp. design, 1925)
Neyman (randomness, sampling, 1925)

jot: representativeness of sample for pop (same as sample, unsample similar)

IID? representative ✓

LD people not random, as they are not at random but from 1
phone books & 2 club membership lists.

bias: a systematic tendency to over- or under-estimate truth.
selection / a systematic tendency bias / to over- or under-represent relevant subsets of pop. (L.D., oversampled on rich people, income assoc. with political affiliation [rich = more republicans])

response / the way the question is asked can influence the answer

Hawthorne effect: people change their behavior when they know they're being watched
nonresponse) people who choose to respond may differ systematically from those who do not in relevant ways.

L.D. had \( \frac{2,400,000}{10,000,000} \times 100\% = 24\% \) response rate (horrible)