C7C **Sample-Resample** Estimate population in a hard-to-count situation (Fish in a lake)

Tag all subjects in sample1 and release. In sample 2, count ttotal number and those with tags.

	Sample1	Sample2	
Tagged	T1	T2	T1: # tagged in sample1
Sample Tot	N1=T1	N2	T2: # tagged in sample2

N1=Size of sample1 N2=Size of sample2

Assume: Tagged subjects mix randomly in entire population; Sample 2 is random sample

Proportional reasoning:

tagged / # in population = fraction observed in 2nd sample T1/N = T2/N2

in population = # in 1st sample divided by fraction in 2nd sample N = T1*(N2/T2)

C7D **Sensitive-question survey so that those answering honestly are not identifiable.**

Group membership is random and known only to the individual.

Those in group 1 agree to say Yes regardless of whether that is true. #No in Group1 = 0

Those who find themselves in group 2 agree to answer honestly. (Concealed by group1)

Subjects	Group1	Group2	
Yes	N/2	=#Yes-N/2	#Yes
No	0	=#No	#No
	N/2	N/2	N

N = Number in both groups combined

N Can be any number, since answer is a ratio.

1. Observe & count Yes and No
2. Infer Group1=Group2 = N/2
   Assume randomly assigned to gr
3. Fraction of Yes who are real
   = (#Yes-N/2)/#Yes
4. Same as fraction of population

C7G **Confidence Intervals: Estimate the interval for the population statistics**

Point estimate = Sample mean or sample proportion

CI = sample mean (proportion) $\pm$ Margin of error (sampling variation)

ME = Margin of [sampling] error n is the size of the sample

For measurements, ME = 2*s/sqrt(n) where s is standard deviation

For proportions, ME = 2 sqrt[p(1-p)/n] where p is proportion in sample

Conservative ME (proportion) = 1 / Sqrt(n) Assumes that p = 1/2

Note: Don't mix percentages and decimal fractions. Use the latter!!!

C7I **Required Sample Size so Margin of Error is Less than M**

For measurements $n = (2*s/M)^2$

For proportions $n = [4*p*(1-p)]/M^2$

$n = 1/M^2$ using conservative estimate.

C7J & C7L **Statistical Significance: Difference in population statistics is likely to be real (not chance).**

* If confidence intervals do not overlap or touch, difference is statistically significant
* If confidence intervals overlap or touch, difference is NOT statistically significant

Note: Statistically significant does not mean "important"

Statistically-significant means difference is statistically unlikely (if due to chance)

1
2

oups