AMS7
Lecture 2
6/22/2016

n = sample size of subjects sum = #
N = total mean = proporte
binary: 1/0 yes/no (binary type of dichotomous)
dichotomous: any other 2 variables

parameter: number that summarizes boundry

An estimate $\hat{\theta}$ of a parameter θ = statistic

When selecting sample, random $\xrightarrow{\text{promotes}}$ similarity
→ mean of sample $\hat{p}$ and p in pop.

Random sampling: Jerzy Neyman 1927

$y \rightarrow$ common variable

$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \text{ or } y_{201} \end{bmatrix}$ n=201

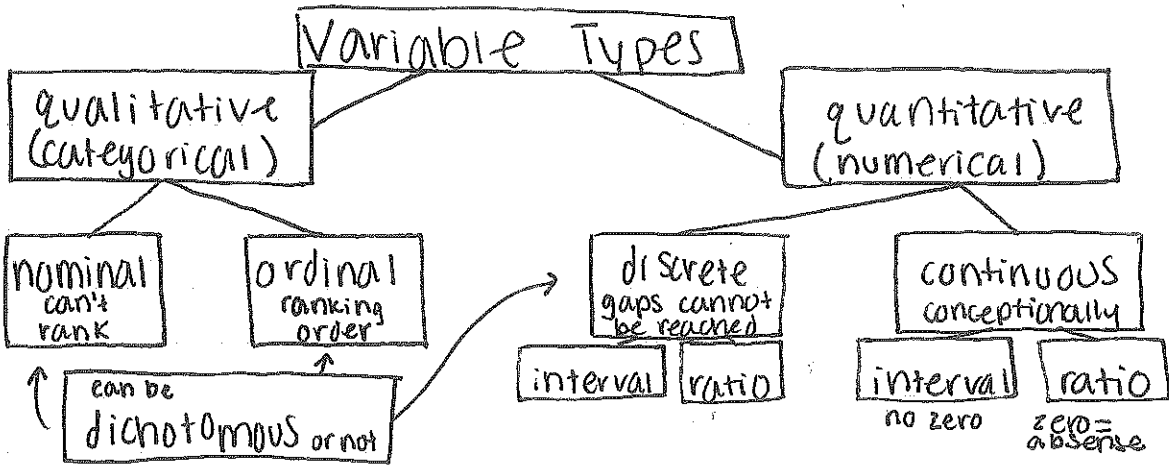
mean: $\bar{y} = \frac{y_1 + y_2 \cdots + y_n}{n} = \frac{1}{n}(y_1 + y_2 + \cdots y_n)$

Sum Notation: $\frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$

i = index of summation

• sum as i runs 1 → n

Two ways for random sampling:
① independent identically distributed (IID)
at random w/ replacement
② simple random sampling (SRS)
at random w/o replacement
more informative, harder math



Note: · truly continuous can never be attained
· when using dich binary coding it goes to qual
· time is interval

Variable	Values	Type
eye color	brown, blue	qual, dichotom.
success @ running maze	1 (very slow), 2 (slow), 3(mod), 4 (fast), 5 (very fast)	qual, ordinal
plant: height	63.2cm, 61.8cm	quan, cont, ratio
# leaves	47, 62	quan, discrete, ratio
growing T for most buds	81.4°F, 79.7°F	quan, cont, interval

1.3

Graphical Descriptive Methods

ex. Butterfly Wing Length

sample (observed butterflies)

(Y) wing length (cm)

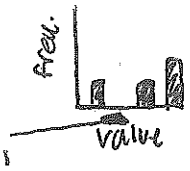
$$\begin{bmatrix} y_1 = 4.4 \\ y_2 = 3.6 \\ \vdots \\ y_n = 3.9 \end{bmatrix} \quad n = 24 \qquad \bar{y} =$$

sort data (number smallest → largest values)

value	freq. (raw freq) / count
3.3	1
3.4	0
3.5	1
3.6	2

$\Sigma = n$

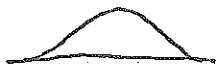
Distribution (raw frequency)



← histogram only for quant.

mode = highest freq.

(if qual, bar or graph)



bell curve
normal curve

Histogram bars must touch (or rep absense)