

Lecture 5

ITRN 501 Fall 2008

Pelletiere

Counts, shares, ratios, measures of
central tendency, spread,
distribution.

Homework 1

- Most everyone received full points
 - Points were deducted if you did not make any attempt to tell me about the issues you encountered in preparing the data.
- Full points just says satisfactory, on the exam there is a higher standard so look at comments.
- Do not provide more data than is requested/relevant.
- Issues of precisions and units.
 - What is the precision of measurement?
 - What is the relevant level of precision?
- Full titles, sources, column headings, notes on each table.
- The initially easiest format or source can be problematic later on.

Descriptive statistics

- Descriptive or summary statistics help us organize, compare and understand data.
- Causal statistics help us determine whether there is reason to believe one thing causes another.
 - Descriptive statistics do not.

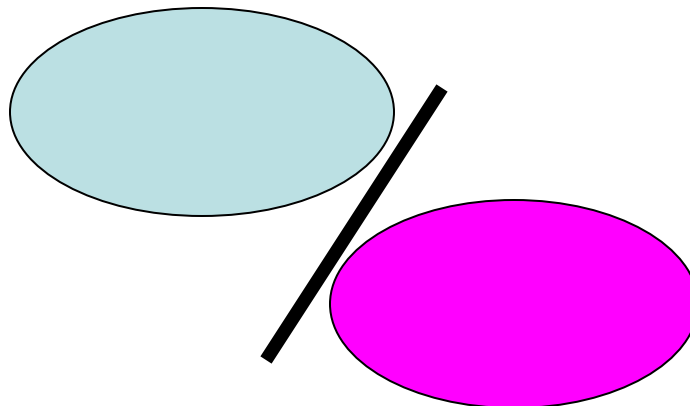
Sum

- Sum of values
- Sum of observations
 - In 2006, with 15 states [observations] participating trade totaled \$13 billion dollars [sum of values]. In 2007, with 32 states participating, the value of trade jumped to \$123 billion.

Ratio

- One series of data compared to a *different* series of data.
 - E.g. Price to earnings ratio, health care expenditures per capita, dollar value of imports per dollar value of exports or even a statement such as “Ghana imports two barrels of oil for every one that Senegal imports.”
 - Often presented as 1:2 rather than as $\frac{1}{2}$ or .5.

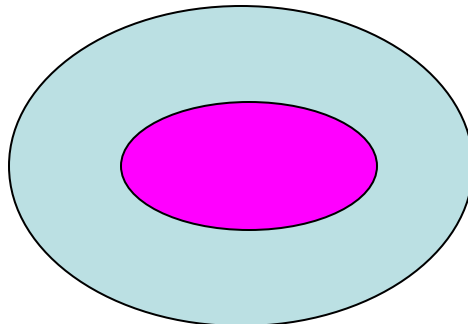
Schematically →



Proportions

- A proportion is the share that a sub or component population makes up of a larger total population
 - Manufacturing exports as a proportion of all exports.
 - Expressed in a per unit basis.
 - $\$100 \text{ million} / \$200 \text{ million} = \$0.50$ of manufacturing trade makes up every dollar of trade.
 - A **percentage** is a proportion multiplied by 100
 - Expresses on a “Per 100” basis
 - 50% of the value of exports is in manufactured goods, i.e. \$50 of every \$100 of exports are manufactured goods.

Schematically →



Frequency distribution (with ordinal data)

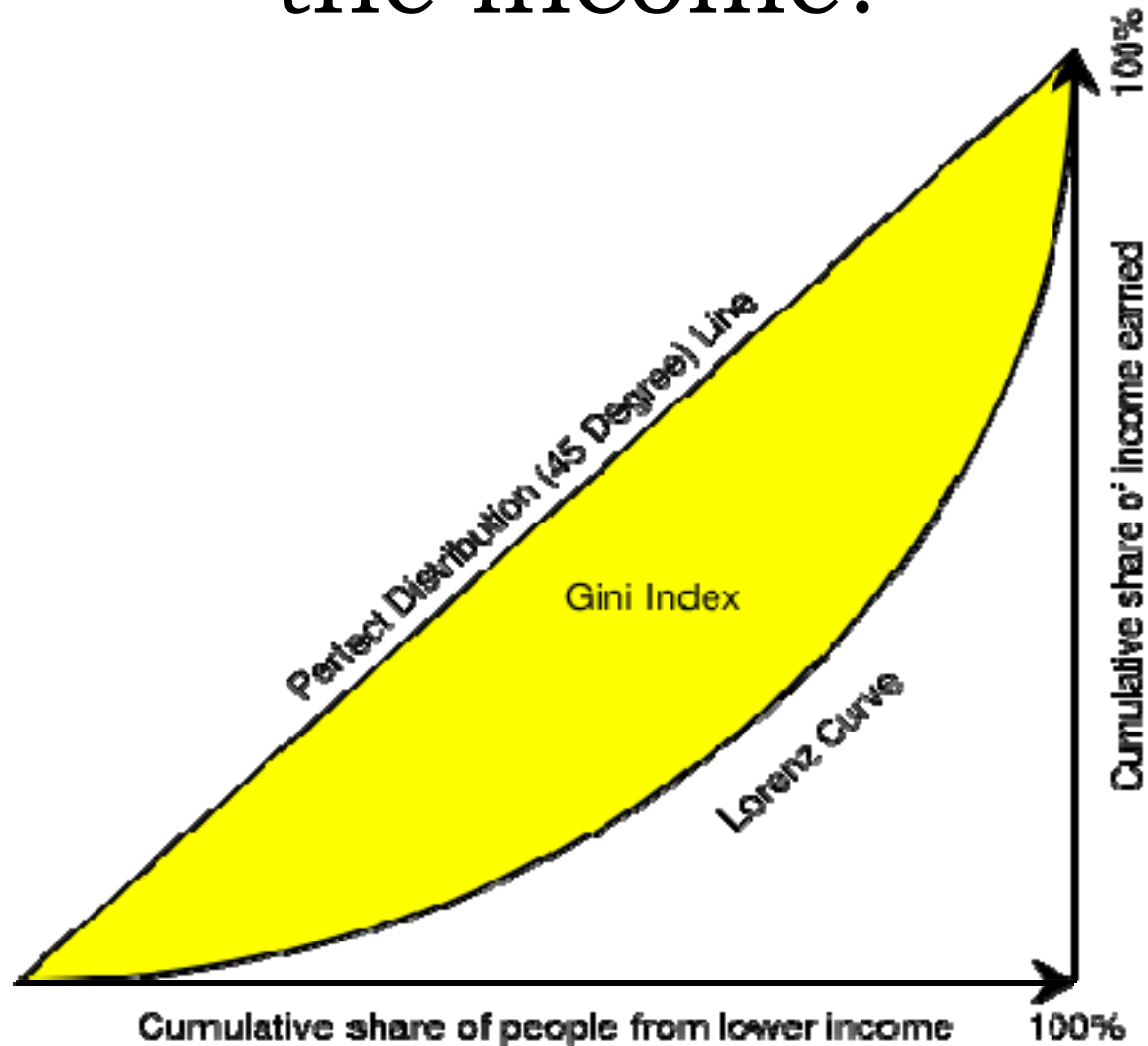
Table 1. An Ordinal Measure of Used Automobile Protection

Score	Policy Summaries	Number of Countries (Total = 132)	Percent
0	No additional restrictions on imports or 'Blue Book' valuation applied.	58	44
1	Taxes escalate with the age of the vehicle; capped depreciation; age limit of 6 years or older applied; or a small additional fee/duty.	24	18
2	Age limits of 5 years or fewer applied; cannot be imported fully assembled; or a combination of these or lesser restrictions.	29	22
3	Imports prohibited; required import licenses not being approved.	21	16

Shares and frequency distributions

- You can look at different shares of the population.
 - The 20% of the population with the lowest incomes
 - The top 10% of trading partners by volume.
- You can compare various “percentiles” to each other.

What percentile of the Population has what share of the income?



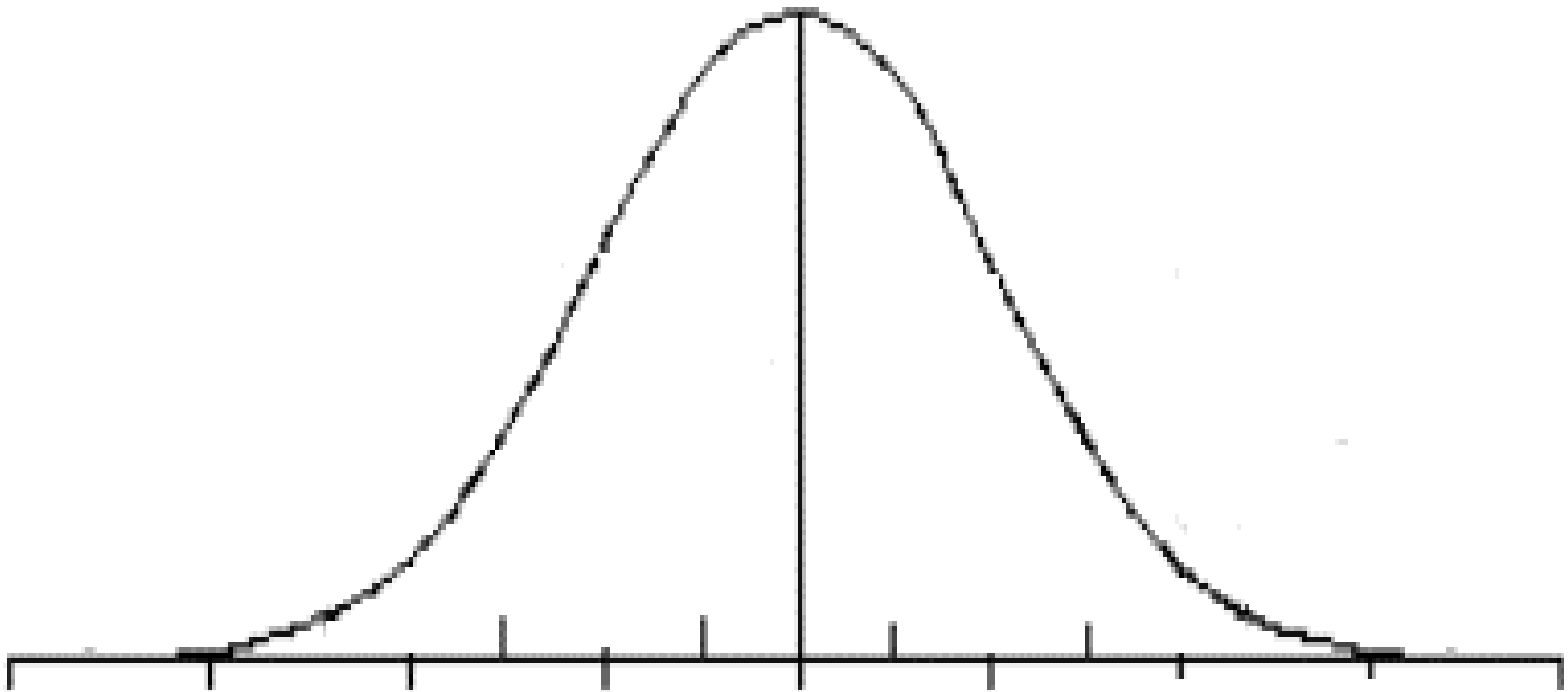
Measures of central tendency

- Mean
 - Sum of all values / number of values
 - The “balancing point” of the data
- Median
 - The middle value when data are arranged in numerical value, from low to high
 - When there is no middle it is the average of the two middle values.
- Mode
 - Single most common value

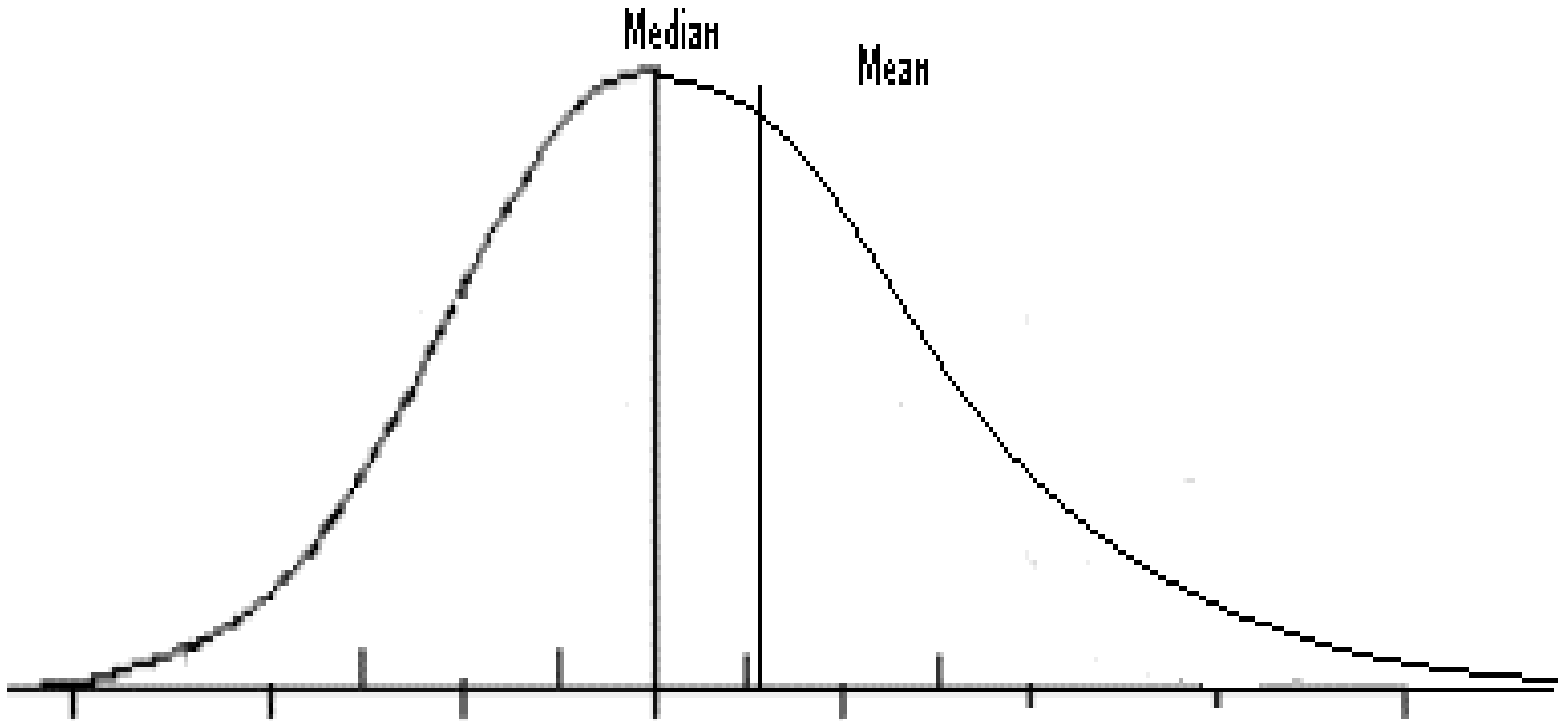
Median

- Different from mean in that it tells us something about the distribution of the value within the population
 - Median is the 50th percentile, immediately we know that half of the units fall below this value and half above.
 - The Nth percentile...
 - Values at other percentiles may also be of interest: quartiles, quintiles etc.

The Normal Curve



Using the mean and median to determine skew



An exercise

- Tell me the date (M/yr) you last moved.
- Organize in graph/timeline
- Determine mean (dates as continuous)
- Determine median
- Determine range
- Determine quartiles

Weighted means

- Allows you to combine means of subsamples to get the overall sample mean when you do not have access to the microdata.
- Multiply values by “weight” and divide by the sum of the weights.
 - $(\text{Value1} * \text{weight1}) + (\text{Value2} * \text{weight2}) \dots + (\text{ValueN} * \text{WeightN}) / (\text{weight1}) + (\text{weight2}) \dots + (\text{WeightN})$

Gaske practice

- Do percentage application pg. 15
- Do mean application pg. 27
- Do median application pg. 28