Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Capacity Building – Module 5

Introduction to Survey Data Analysis & Use

Elke de Buhr, PhD
Payson Center for International Development
Tulane University

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Statistics

- A field of study concerned with
  - The collection, organization, summarization, and analysis of data
  - The drawing of inferences about a body of data when only a part of the data is observed

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Type of Statistics

- Descriptive data analysis
  - Organizing and summarizing data
- Statistical inference
  - Procedure by which we reach a conclusion about a population on the basis of the information contained in a sample that has been drawn from that population

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Monitoring Strategy

- Comparison
  - Same group
  - Different groups
- Design balances accuracy and reliability with cost and feasibility

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Sample Selection

- Sample size
- Sampling frame
- Sample selection = sampling
  - Probability sampling
  - Nonprobability sampling

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Data Collection Methods

- Quantitative
  - Record reviews
  - Formal Surveys
  - Others
- Qualitative
  - Open ended interviews (key informants, etc.)
  - Focus group discussions (FGD)
  - Matrix ranking (preference)
  - Direct Observations
  - Others

(IDRC, p. 149)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Differences between Methods

- Quantitative methods
  - Quantification in collection and analysis of data
  - Testing of theories, deductive
  - Incorporates natural, scientific model
  - Views social reality as an external, objective reality
- Qualitative methods
  - Qualification of words/narrative in collection and analysis of data
  - Generation of theories, inducive
  - Emphasizes ways in which individuals interpret their social world
  - Views social reality as a constantly shifting property of individuals' creation

(Bryman, Alan **Social Research Methods**, 2nd Ed. Oxford: Oxford University Press, 2004, p. 20-21)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Concepts, Variables and Indicators

	Example 1	Example 2	Example 3
Concepts	Size	Economic well-being	Health
Variables	Area	Income per capita	Life Expectancy
Indicators	Square kilometers	Purchasing Power Parity (PPP) GNP ($) per capita	Average years of life if born in 1970

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Categorical vs. Continuous Variables

- Continuous variables
  - A variable that can be measured (weight, height, age, etc.)
- Categorical variables
  - A variable that cannot be measured but can be categorized (ethnic group, age group, educational level, socio-economic class, etc.)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Categorical Scale

- Nominal scale
  - Lowest measurement scale, consists of "naming" observations or classifying them into various mutually exclusive and collectively exhaustive categories
  - Any assigned numerical value is merely for convenience (e.g. Christian = 1, Jewish = 2, Buddhist = 3)
- Ordinal scale
  - Observations are not only different from category but can be ranked according to some criterion, they are said to be measured on an ordinal scale
  - Assigned numerical value reflects rank order (e.g. Low socioeconomic status = 1, Medium socioeconomic status = 2, High socioeconomic status = 3)

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Continuous Scale

- Interval scale
  - Measurement that has order and the distant between any two measurements is known
  - No true zero point (e.g. temperature)
- Ratio scale
  - Characterized by the fact that equality of ratios as well as equality of intervals may be determined
  - True zero point, zero indicates absence (e.g. height, length and weight)

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Level of Measurements

	Mutually Exclusive (Distinction of difference)	Ordered (Distinction of the direction of difference)	Equal Intervals (Distinction of amount of difference)	True Zero Point
Nominal	X			
Ordinal	X	X		
Interval	X	X	X	
Ratio	X	X	X	X

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Categorical Scale

- Nominal scale
  - Lowest measurement scale, consists of "naming" observations or classifying them into various mutually exclusive and collectively exhaustive categories
  - Any assigned numerical value is merely for convenience (e.g. Christian = 1, Jewish = 2, Buddhist = 3)
- Ordinal scale
  - Observations are not only different from category but can be ranked according to some criterion, they are said to be measured on an ordinal scale
  - Assigned numerical value reflects rank order (e.g. Low socioeconomic status = 1, Medium socioeconomic status = 2, High socioeconomic status = 3)

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Continuous Scale

- Interval scale
  - Measurement that has order and the distant between any two measurements is known
  - No true zero point (e.g. temperature)
- Ratio scale
  - Characterized by the fact that equality of ratios as well as equality of intervals may be determined
  - True zero point, zero indicates absence (e.g. height, length and weight)

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Data Analysis - Quantitative

- Type of variable
  - Categorical
  - Continuous
- Descriptive analysis
- Hypothesis testing

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Descriptive Data Analysis

- Categorical/Continous data
  - Frequency tables
- Continuous data
  - Central tendency
  - Variability

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Example of Frequency Table for a Nominal-Level Variable

Table 1: Developing Countries by Region

Region	Frequency (*f*)
East Asia/Pacific	23
Europe/Central Asia	28
Latin America/Caribbean	33
Middle East/North Africa	16
South Asia	8
Sub-Saharan Africa	49
Total (N)	**157**

(World Bank, World Development Indicators 2002, inside back cover)

Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Example of a Frequency Table for an Ordinal-Level Variable

Table 2: Developing Countries by Income Group

Income Group (GNP/Capita)	Frequency (*f*)	Percentage (%)
Low Income (<$761)	63	40
Lower-Middle Income ($761-3,030)	58	37
Upper-Middle Income ($3,031-9,360)	36	23
Total (N)	**157**	**100**

(World Bank, World Development Indicators 2002, inside front cover)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Example for Ratio Variable

Table 3: Developing Countries by Health Expenditure per Capita

PPP $/capita Most Recent Yr	Freq. (*f*)	Percent. (%)	Cumulative Percent. (%)
$50 or less	22	23	23
$51 to $100	17	18	41
$101 to $300	25	26	67
$301 to $500	17	18	85
$501 to $1000	13	14	99
$1001 or more	1	1	100
Total (N)	**95**	**100**	

(World Bank, World Development Indicators 2000, Table 2.14)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Continuous Data

- Central tendency
  - Average or mean
  - Median
  - Mode
- Variability
  - Range
  - Variance
  - Standard deviation

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Measures of Central Tendency

- Mode
  - Most frequent score
  - Only appropriate measure of central tendency for categorical data without order (nominal scale)
- Median
  - Middle value of a distribution once the values have been ranked (Median of the numbers 1,2,3,4,5 is 3)
  - If the sample contains an even number of observations, the median is the average of the middle two numbers (Median of the numbers 1,2,3,4,5,6 is (3+4)/2 = 3.5

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Measures of Central Tendency (cont.)

- Mean
  - Sum of the observations in the sample or population, divided by the number of observations
  - Arithmetic average (Mean of the integers one through five is (1+2+3+4+5)/5 = 15/5 =3)

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Measures of Central Tendency (cont.)

Measures of Central Tendency	Appropriate Data Type Application
Mode Most frequently occurring value	**Nominal, Ordinal,** and (sometimes) **Interval, and Ratio** Data
Median Exact center of rank-ordered data or average of two middle values	**Ordinal-Level and Ratio and Interval** Data (particularly when skewed)
Mean Arithmetic average	**Ratio or Interval-Level** Data (and, though controversial, some ordinal-level data)

(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Measures of Variability

- Range
  - Largest value minus the smallest value
- Variance
  - Sum of the deviation of each variable from the mean, squared, divided by N or by n-1, where N is the size of a population, and n is the size of a sample
- Standard Deviation
  - Square root of the variance (Standard deviation of the numbers 1,2,3,4,5 is $\sqrt{2.5} = 1.58$)

(Phuong Pham, Introduction to Quantitative Analysis)

- Population Variance: σ^2 $\dfrac{\sum_{i=1}^{N}\left(x_i-\mu\right)^2}{N}$

- Sample Variance: = s^2 $\dfrac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}{n-1}$

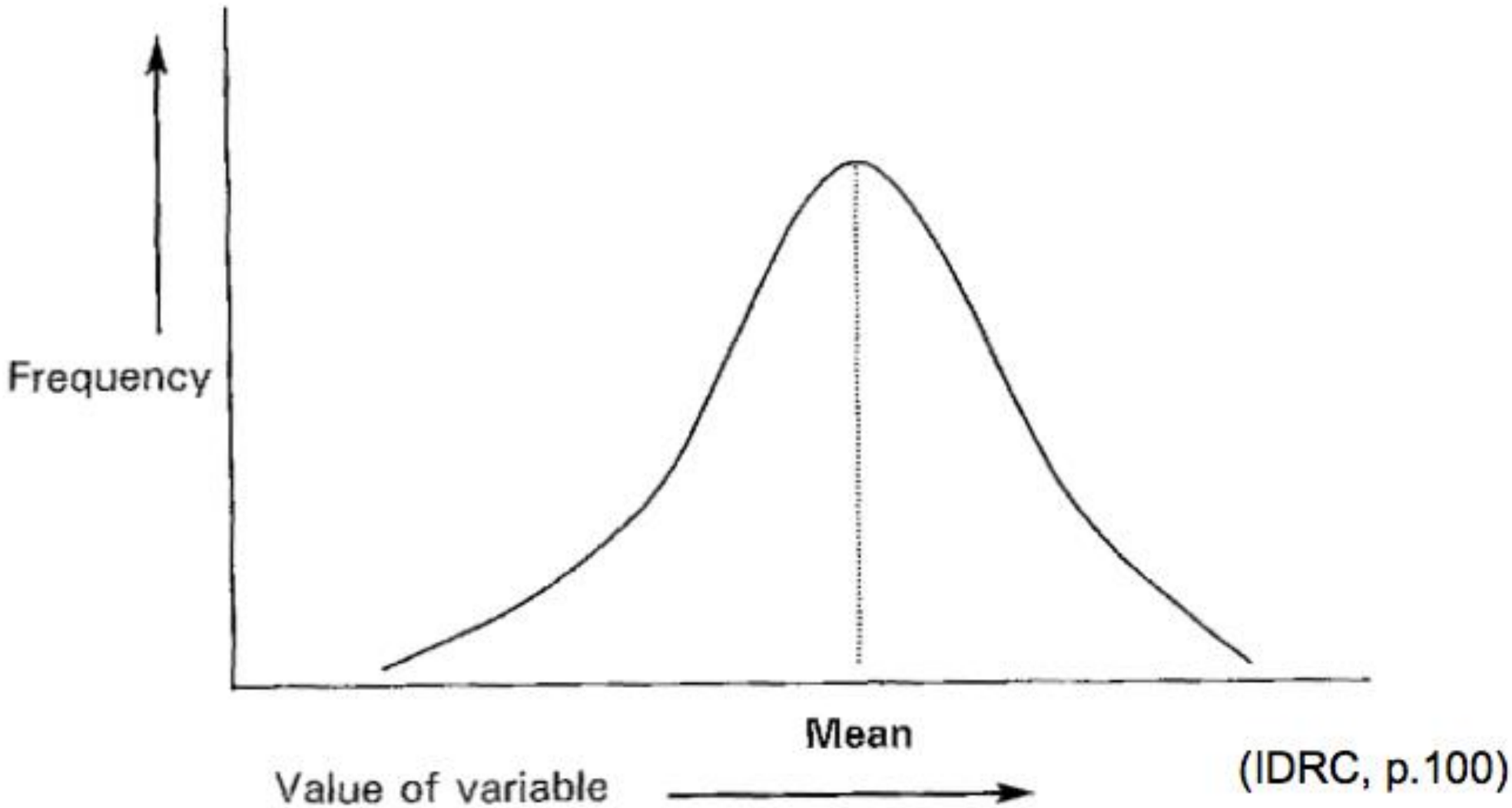
(Phuong Pham, Introduction to Quantitative Analysis)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Normal distribution

Figure 27.1: Normal distribution curve



(IDRC, p.100)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Figures

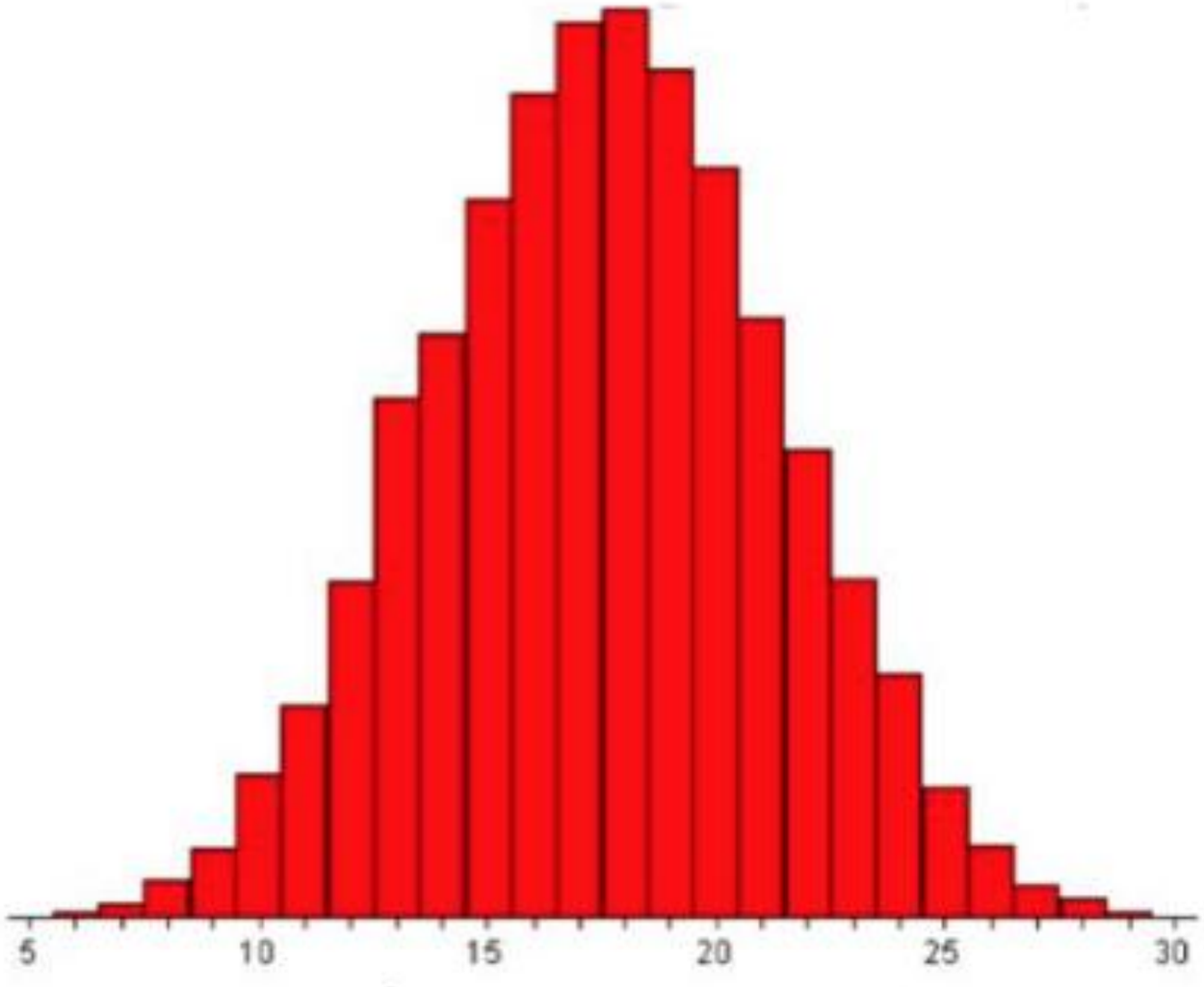
- Categorical data
  - Bar charts
  - Pie charts
- Continous data
  - Histograms
  - Line graphs
  - Scatter diagrams

Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Histogram



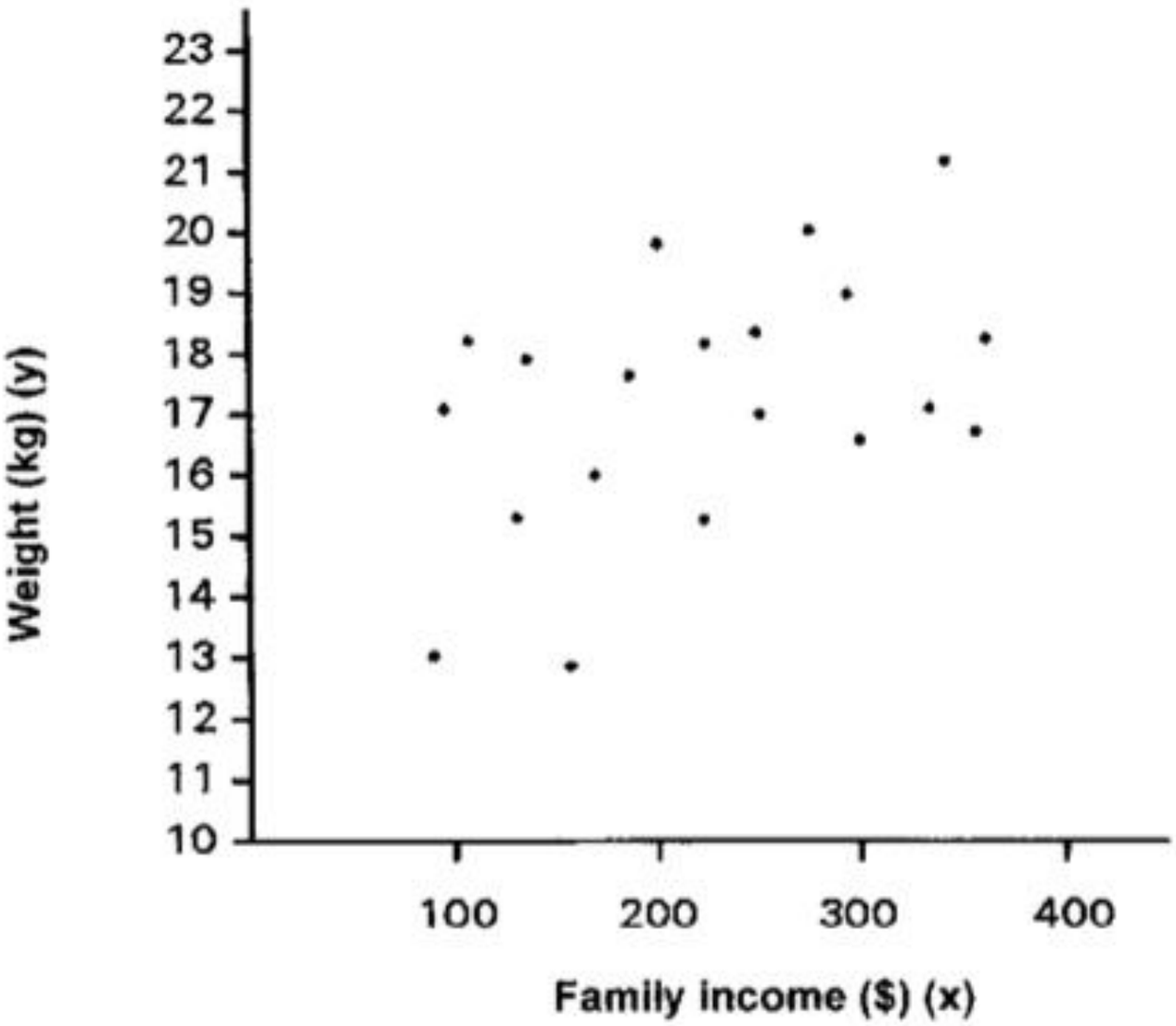
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Scatter Diagram

Figure 31.1: Weights and family incomes of 20 children 5 years of age



(IDRC)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Hypothesis Testing - Quantitative

- Formal statement that predicts relationship between one or more factors and the problem under study.
- Support or reject the null hypothesis
- Null = no relationship
- Test:
  - Compare same variable over time
  - Comparison between 2 or more groups

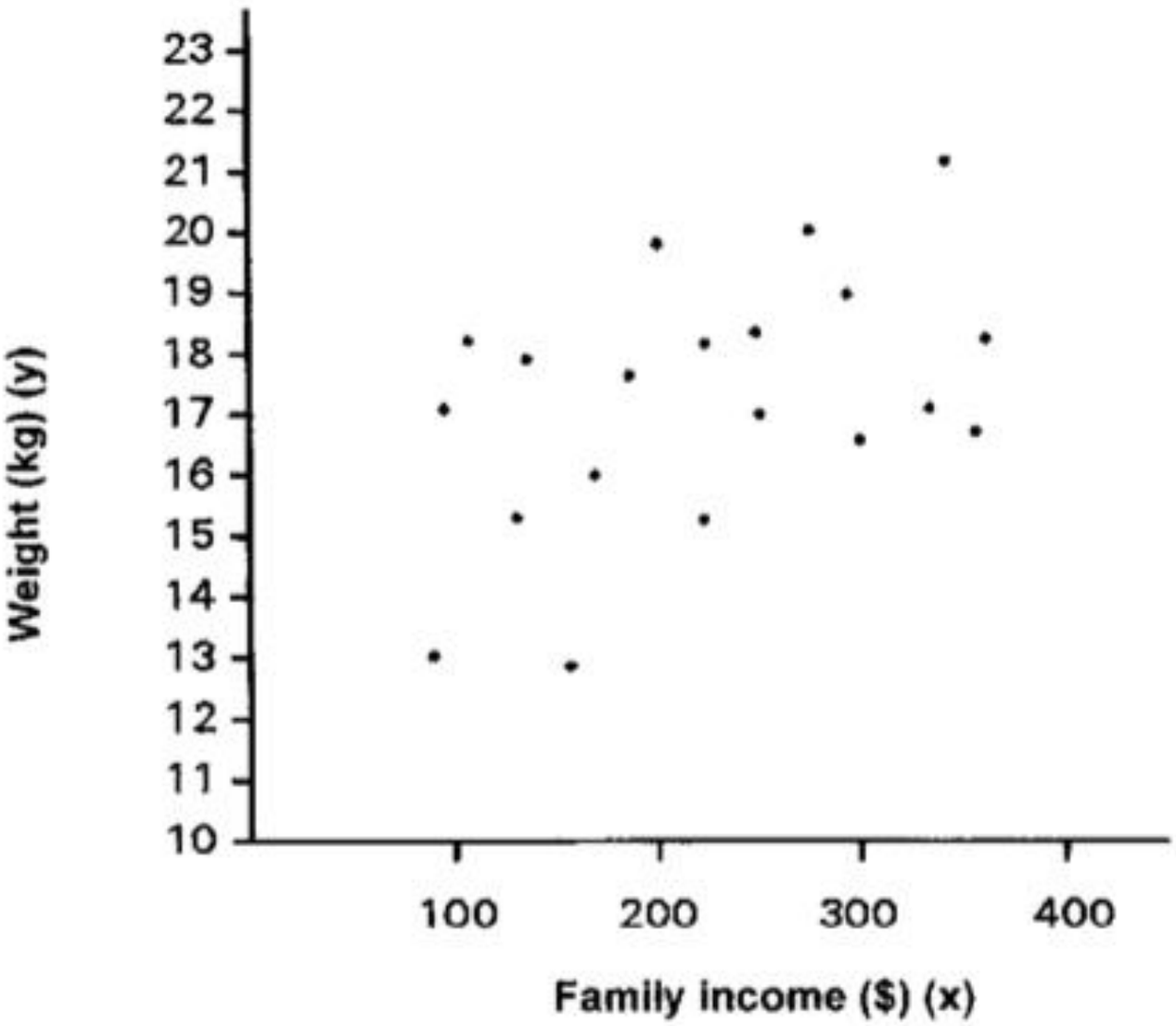
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Scatter Diagram

Figure 31.1: Weights and family incomes of 20 children 5 years of age



(IDRC)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Hypothesis Testing - Quantitative

- Formal statement that predicts relationship between one or more factors and the problem under study.
- Support or reject the null hypothesis
- Null = no relationship
- Test:
  - Compare same variable over time
  - Comparison between 2 or more groups

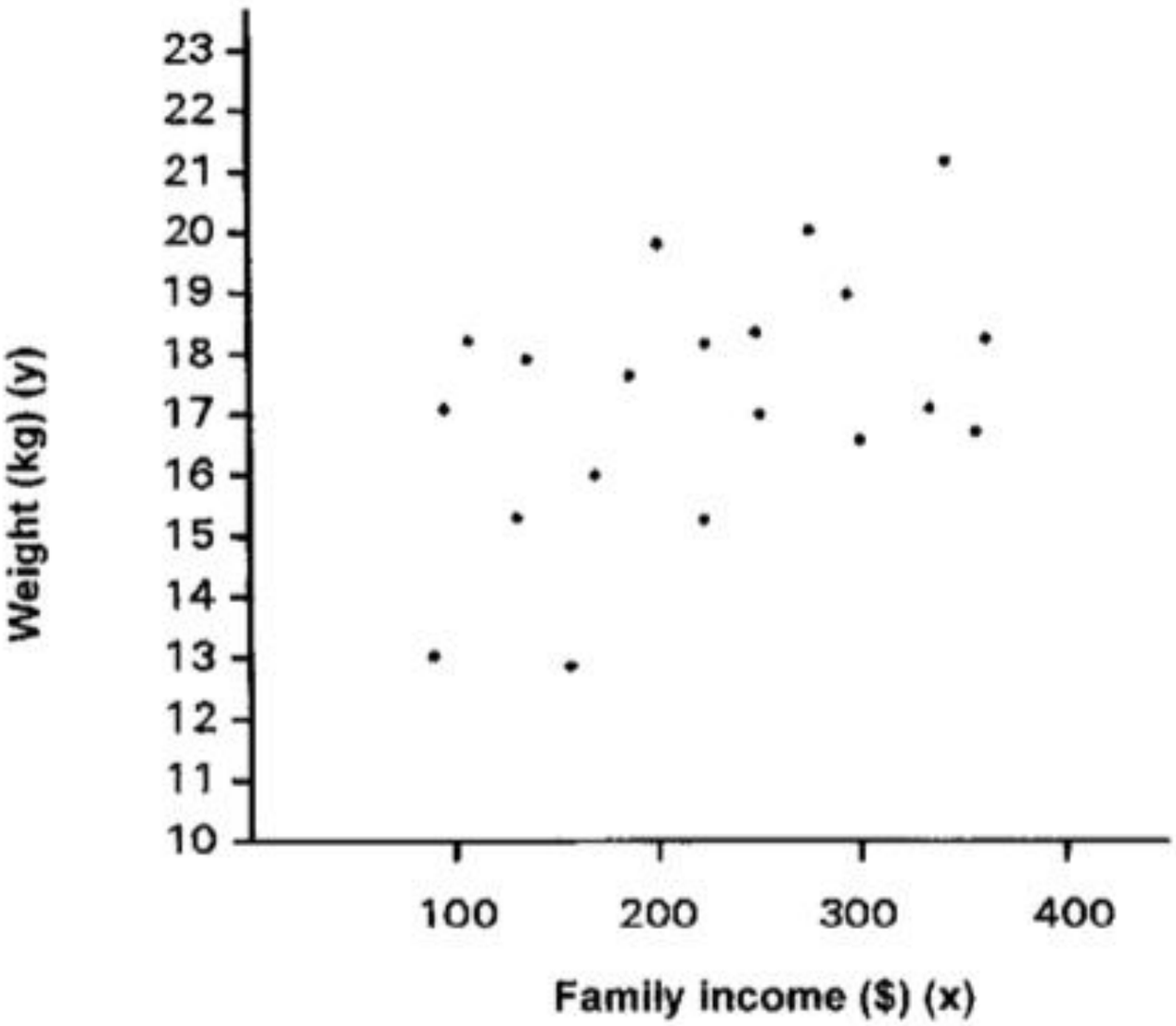
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Scatter Diagram

Figure 31.1: Weights and family incomes of 20 children 5 years of age



(IDRC)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Hypothesis Testing - Quantitative

- Formal statement that predicts relationship between one or more factors and the problem under study.
- Support or reject the null hypothesis
- Null = no relationship
- Test:
  - Compare same variable over time
  - Comparison between 2 or more groups

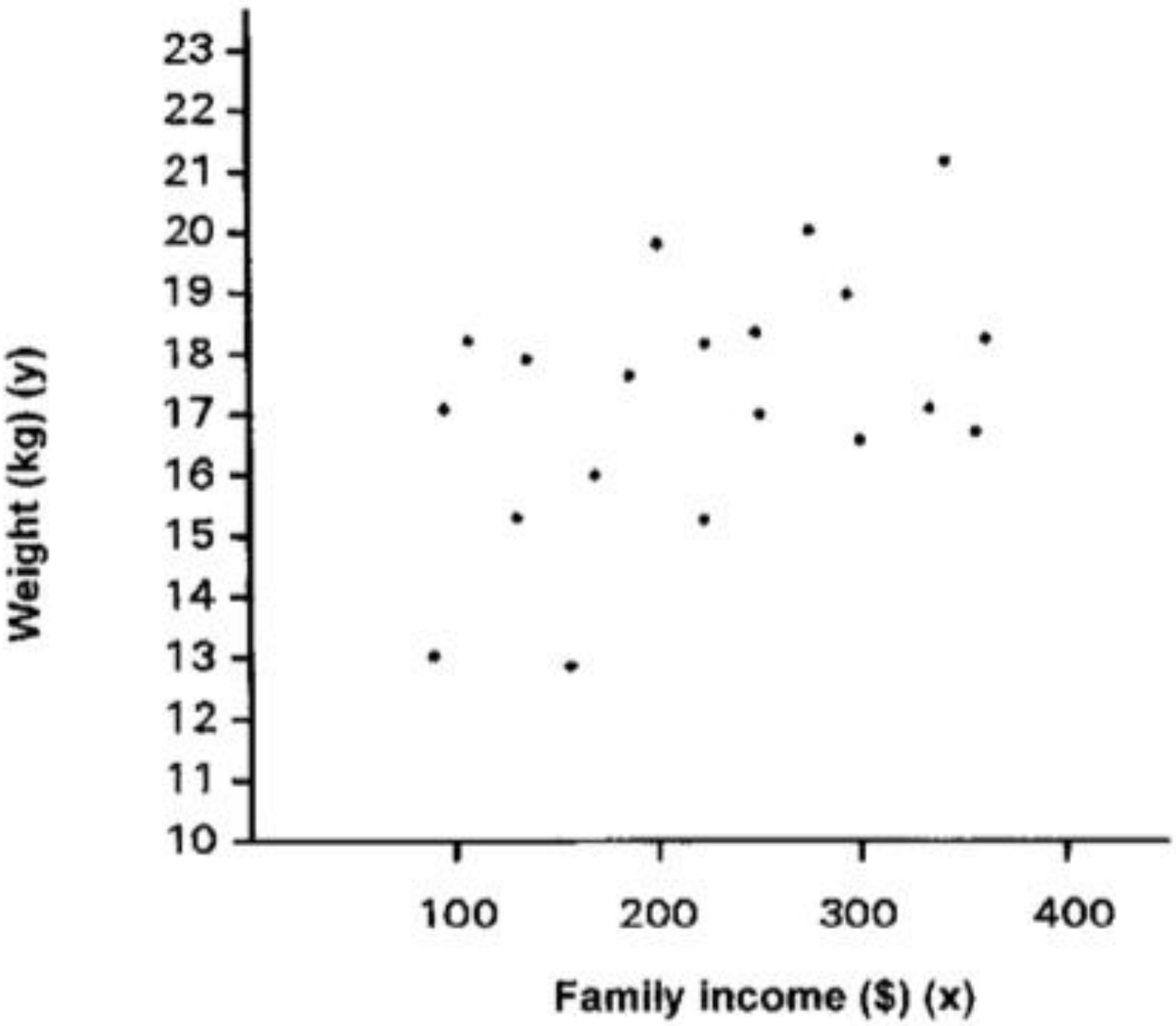
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Scatter Diagram

Figure 31.1: Weights and family incomes of 20 children 5 years of age



(IDRC)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Hypothesis Testing - Quantitative

- Formal statement that predicts relationship between one or more factors and the problem under study.
- Support or reject the null hypothesis
- Null = no relationship
- Test:
  - Compare same variable over time
  - Comparison between 2 or more groups

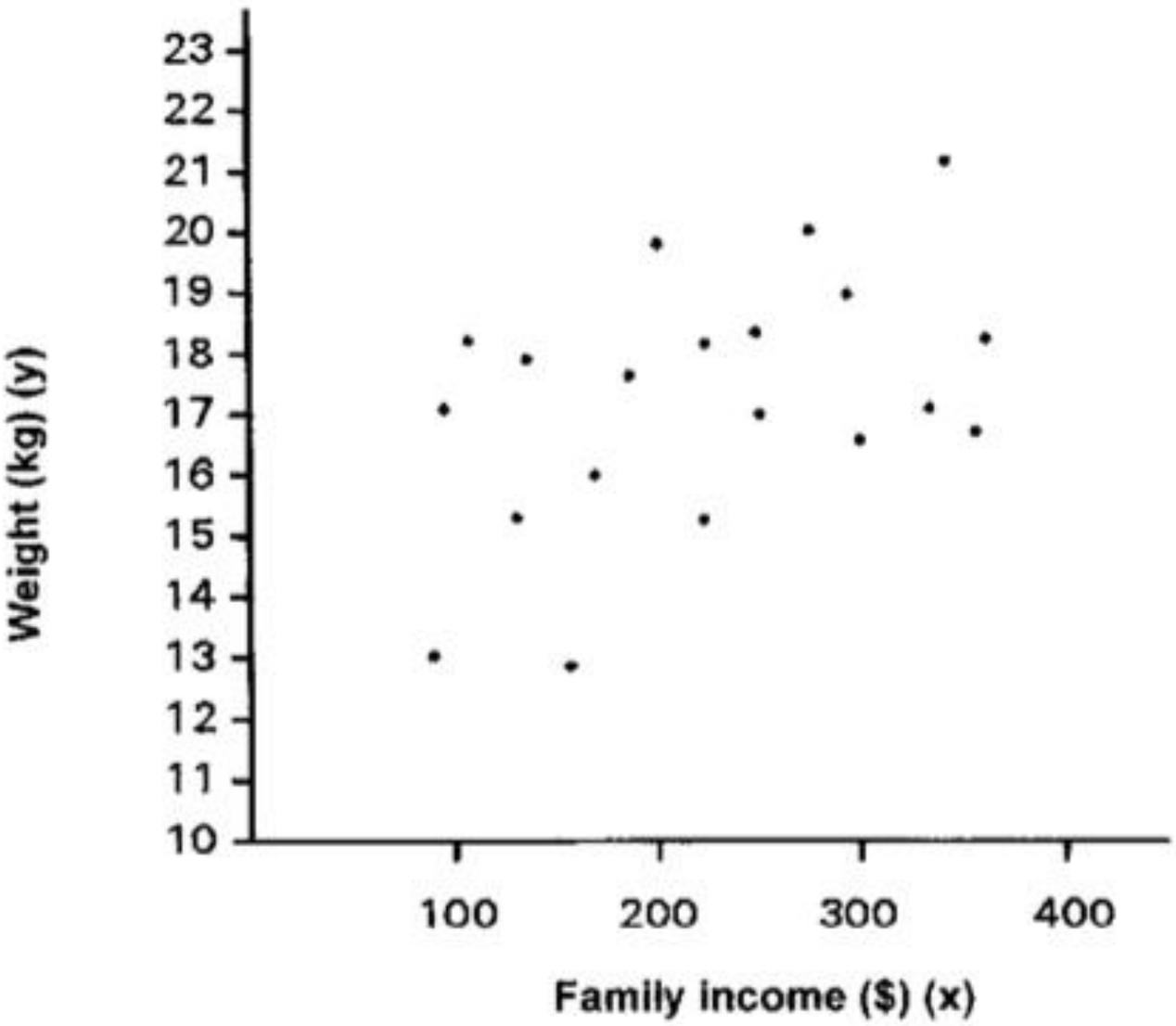
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Scatter Diagram

Figure 31.1: Weights and family incomes of 20 children 5 years of age



(IDRC)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

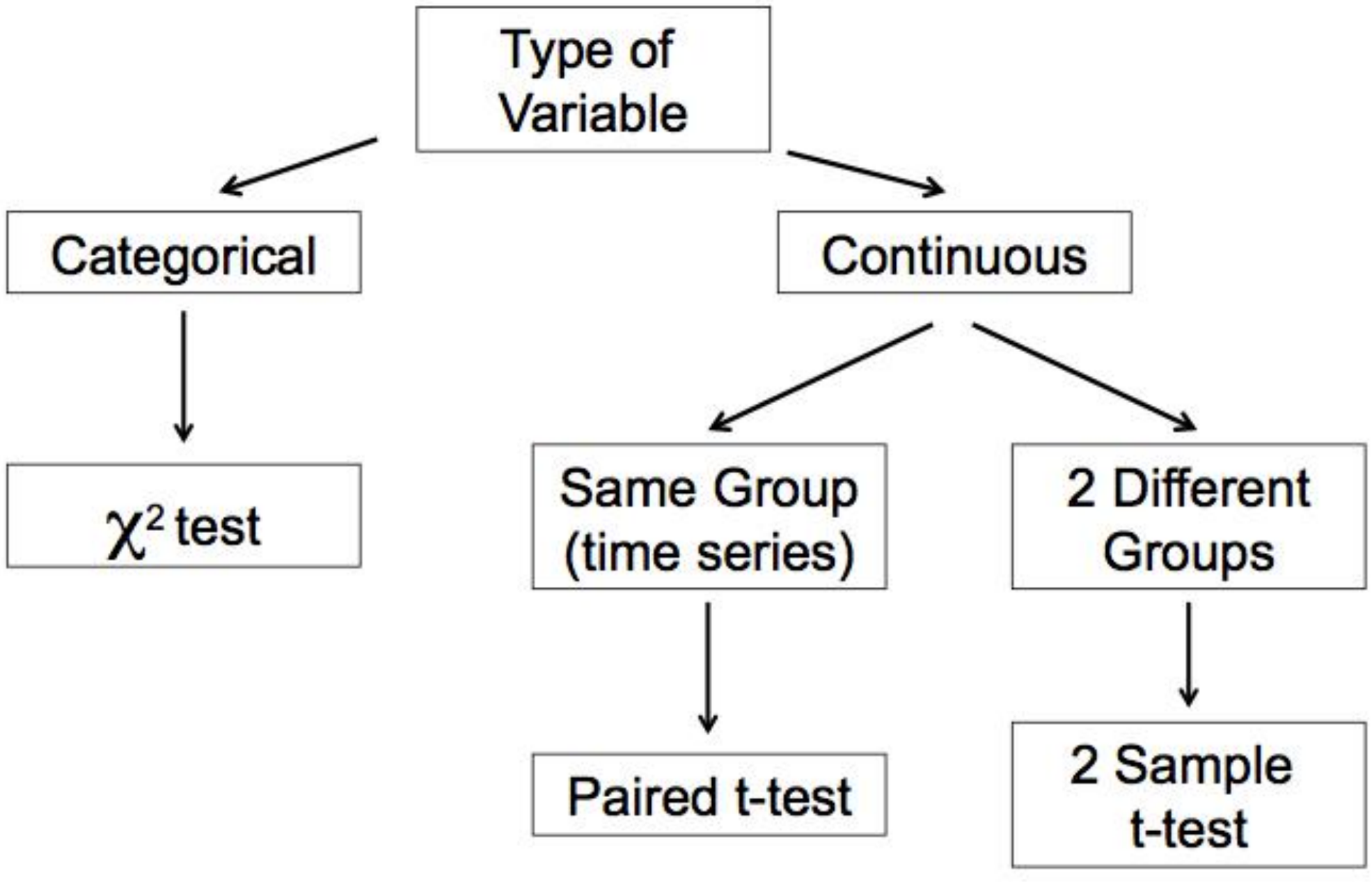
Statistical Inference

1. Develop a hypothesis
2. Formulate the null hypothesis
3. Calculate a test statistic
4. Calculate the probability (p-value) of null hypothesis being false
5. Reject or accept the null hypothesis

Tulane University

Payson

The Payson Center for International Development and Technology Transfer



Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Causal Relationship

- Independent vs. dependent indicators
- Statistical test
- Measures how strong relationship
- Linear correlation coefficient: how well the line drawn fits the observed data points
  - r = 1; perfect fit
  - r = 0; no relationship

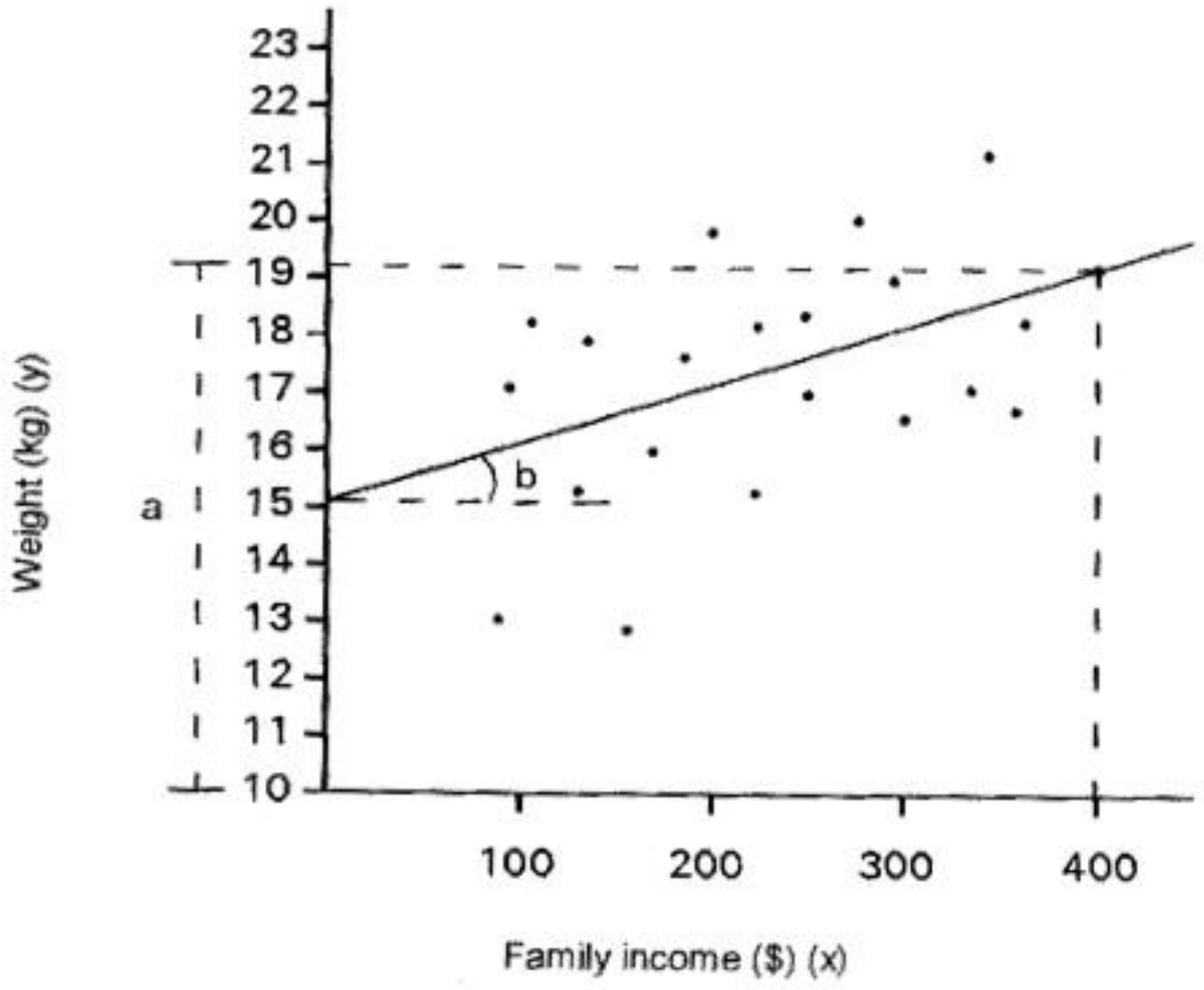
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Linear Regression

Figure 31.2: Linear regression of weight of 5 years old children on family income



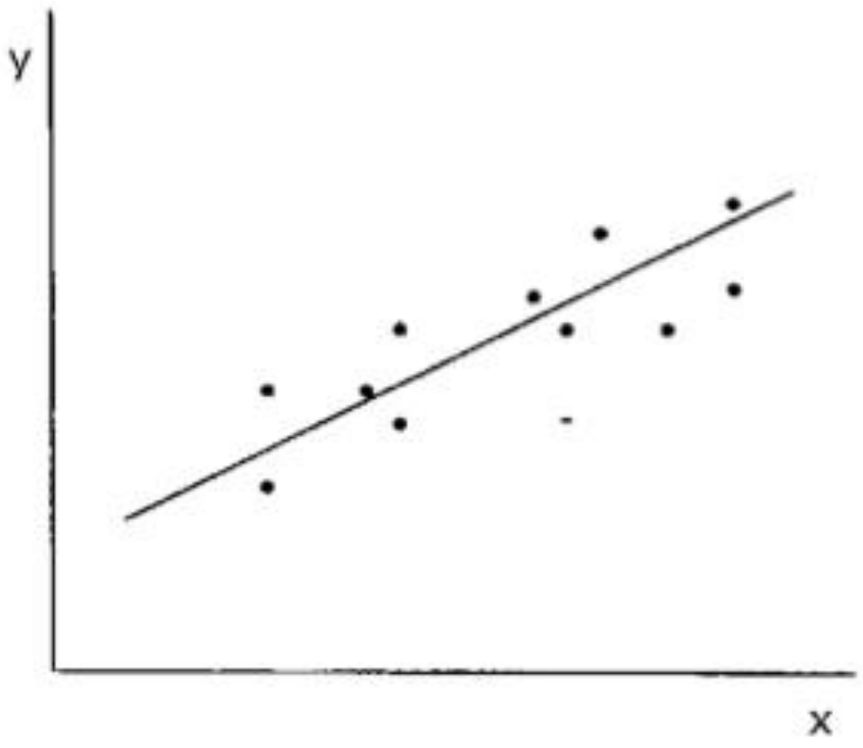
(IDRC, v2)

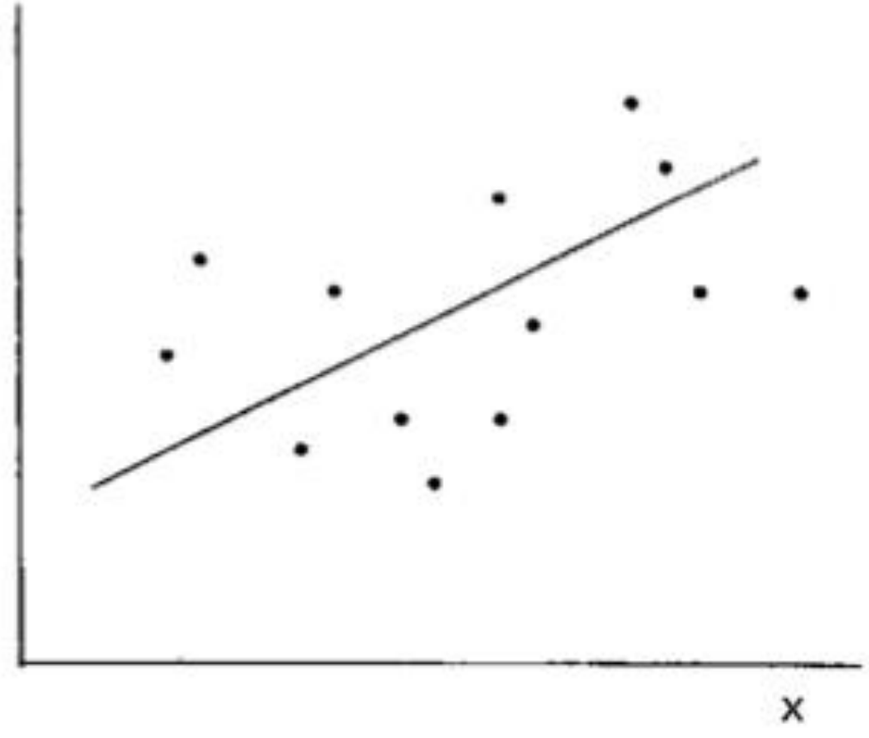
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Relationship





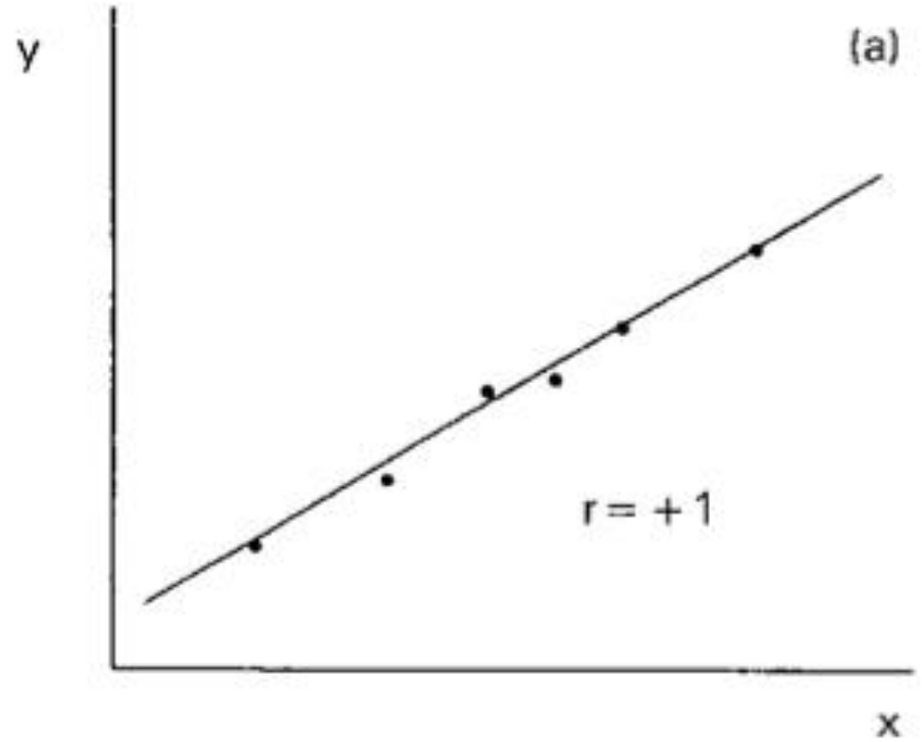
(IDRC, v2)

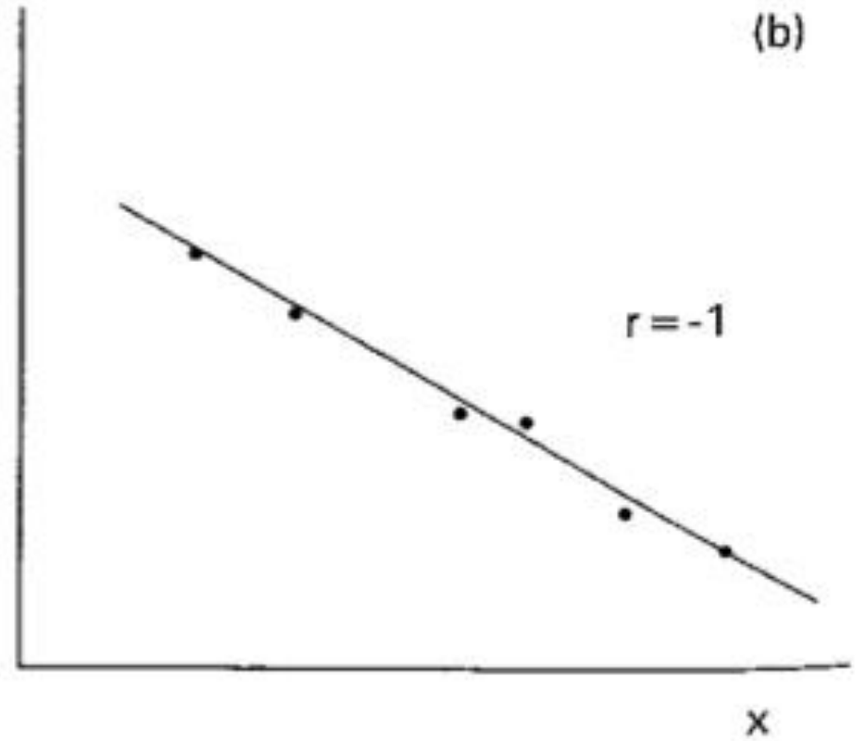
Tulane University

Payson

The Payson Center for International Development and Technology Transfer

Direction of Relationship





(IDRC, v2)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Data Analysis - Qualitative

- Descriptive
  - Classify responses
  - Level of agreement between responses
- Hypothesis testing
  - Comparing responses
  - Relationships between variables

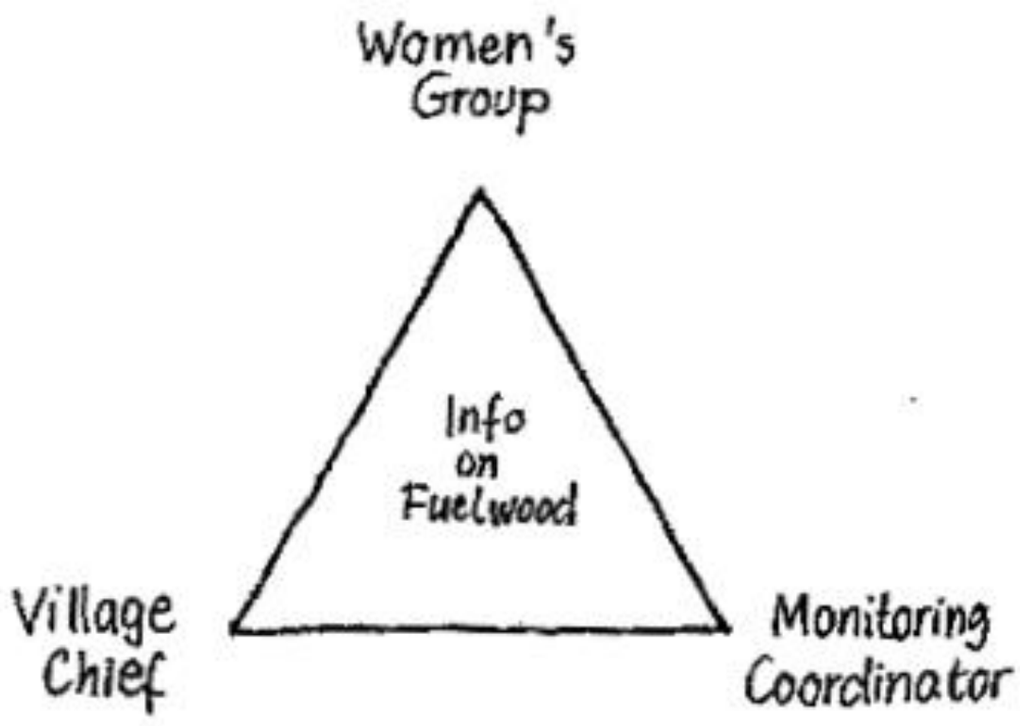
Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Data Analysis - Triangulation

Methods

People Consulted



(Margoluis & Salafsky)

Tulane University

Payson
The Payson Center for International Development and Technology Transfer

Communication of Results

- Often research findings are important to multiple audiences who can be reached by multiple communication channels
- The ideal audience passes the information it receives on to other people (multiplier effect)
- Select format:
  - Type of audience
  - Type of information
  - Cost