Introduction to statistics

Lecture 1

What are statistics

Scenarios

- a new advertisement for Ben and Jerry's ice cream introduced in late May of last year resulted in a 30% increase in ice cream sales for the following three months. Thus, the advertisement was effective.
- The more churches in a city, the more crime there is. Thus, churches lead to crime.
- 75% more interracial marriages are occurring this year than 25 years ago. Thus, our society accepts interracial marriages.

Importance of statistics

- 4 out of 5 dentists recommend Dentine.
- Almost 85% of lung cancers in men and 45% in women are tobacco-related.
- People tend to be more persuasive when they look others directly in the eye and speak loudly and quickly.
- Women make 75 cents to every dollar a man makes when they work the same job.
- A surprising new study shows that eating egg whites can increase one's life span.

"The coming century is surely the century of data."

– David Donoho (statistics professor at Stanford University), 2000

"There are three kinds of lies — lies, damned lies, and statistics

– Mark Twain

Descriptive statistics

Numbers that are used to summarize and describe data



Average salaries for various occupations in 2003

€ 99,540	Pediatricians
€ 95,300	Dentists
€ 92,050	Podiatrists
€ 60,700	Physicisit
€ 45,980	Architects
€ 42,460	Clinical psychologist
€ 40,320	Flight attendants
€ 37,500	Elementary school teachers
€ 35,800	Police officers
€ 21,920	Floral designers

Winning Olympic marathon times

Women					
Year	Winner	Country	Time		
1984	Joan Benoit	USA	2:24:52		
1988	Rosa Mota	POR	2:25:40		
1992	Valentina Yegorova	UT	2:32:41		
1996	Fatuma Roba	ETH	2:26:05		
2000	Naoko Takahashi	JPN	2:23:14		
2004	Mizuki Noguchi	JPN	2:26:20		
Men					
Year	Winner	Country	Time		
1896	Spiridon Louis	GRE	2:58:50		
1900	Michel Theato	FRA	2:59:45		
1904	Thomas Hicks	USA	3:28:53		
1906	Billy Sherring	CAN	2:51:23		
1908	Johnny Hayes	USA	2:55:18		
10.10		A 14			

1920	Hannes Kolehmainen	FIN	2:32:35
1924	Albin Stenroos	FIN	2:41:22
1928	Boughra El Ouafi	FRA	2:32:57
1932	Juan Carlos Zabala	ARG	2:31:36
1936	Sohn Kee-Chung	JPN	2:29:19
1948	Delfo Cabrera	ARG	2:34:51
1952	Emil Ztopek	CZE	2:23:03
1956	Alain Mimoun	FRA	2:25:00
1960	Abebe Bikila	ETH	2:15:16
1964	Abebe Bikila	ETH	2:12:11
1968	Mamo Wolde	ETH	2:20:26
1972	Frank Shorter	USA	2:12:19
1976	Waldemar Cierpinski	E.Ger	2:09:55
1980	Waldemar Cierpinski	E.Ger	2:11:03
1984	Carlos Lopes	POR	2:09:21
1988	Gelindo Bordin	ITA	2:10:32

Inferential statistics

Populations and samples

Sample - small subset of a larger set of data

Population - the larger set of sample data



example

 A substitute teacher wants to know how students in the class did on their last test. The teacher asks the 10 students sitting in the front row to state their latest test score. He concludes from their report that the class did extremely well. What is the sample? What is the population? Can you identify any problems with choosing the sample in the way that the teacher did?

Inferential statistics

Simple Random Sampling

Simple Random Sampling



example

• A research scientist is interested in studying the experiences of twins raised together versus those raised apart. She obtains a list of twins from the National Twin Registry, and selects two subsets of individuals for her study. First, she chooses all those in the registry whose last name begins with Z. Then she turns to all those whose last name begins with B. Because there are so many names that start with B, however, our researcher decides to incorporate only every other name into her sample. Finally, she mails out a survey and compares characteristics of twins raised apart versus together.

Variables

Independent & Dependent



example

 Can blueberries slow down aging? A study indicates that antioxidants found in blueberries may slow down the process of aging. In this study, 19-month-old rats (equivalent to 60year-old humans) were fed either their standard diet or a diet supplemented by either blueberry, strawberry, or spinach powder. After eight weeks, the rats were given memory and motor skills tests. Although all supplemented rats showed improvement, those supplemented with blueberry powder showed the most notable improvement.

• the independent variable is dietary supplement (none, blueberry, strawberry, spinach)

Variables

Discrete & Continuous



Variables

Qualitative & Quantitative



Levels of measurement

- Nominal scales
- Ordinal scales
- Interval scales
- Ratio scales

the end