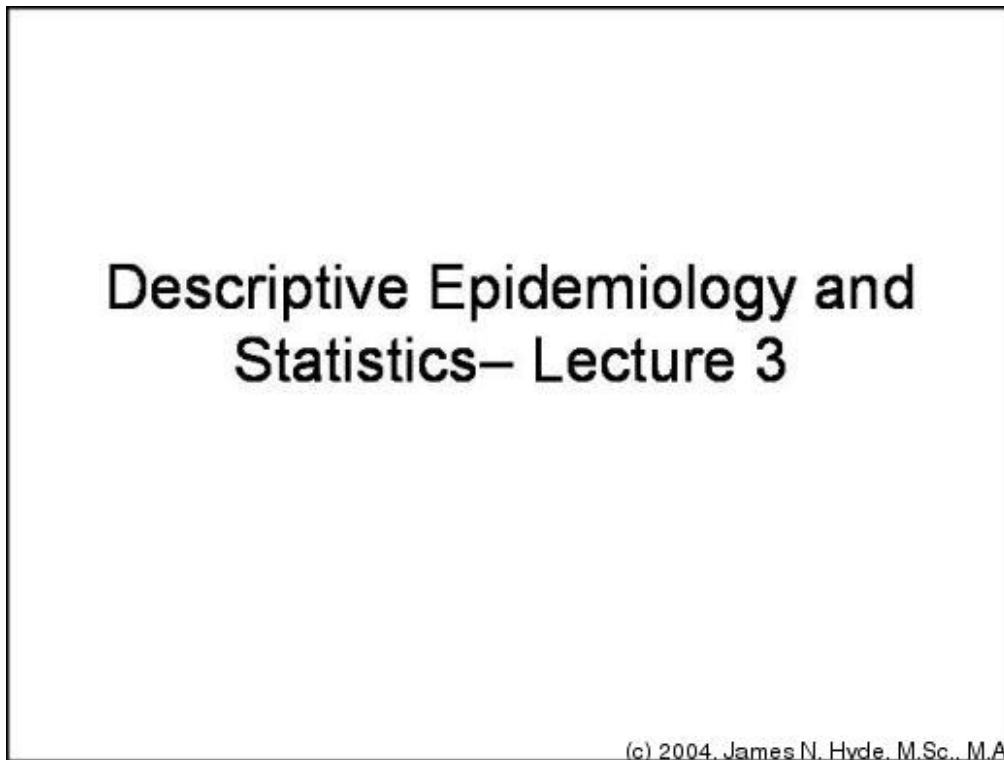
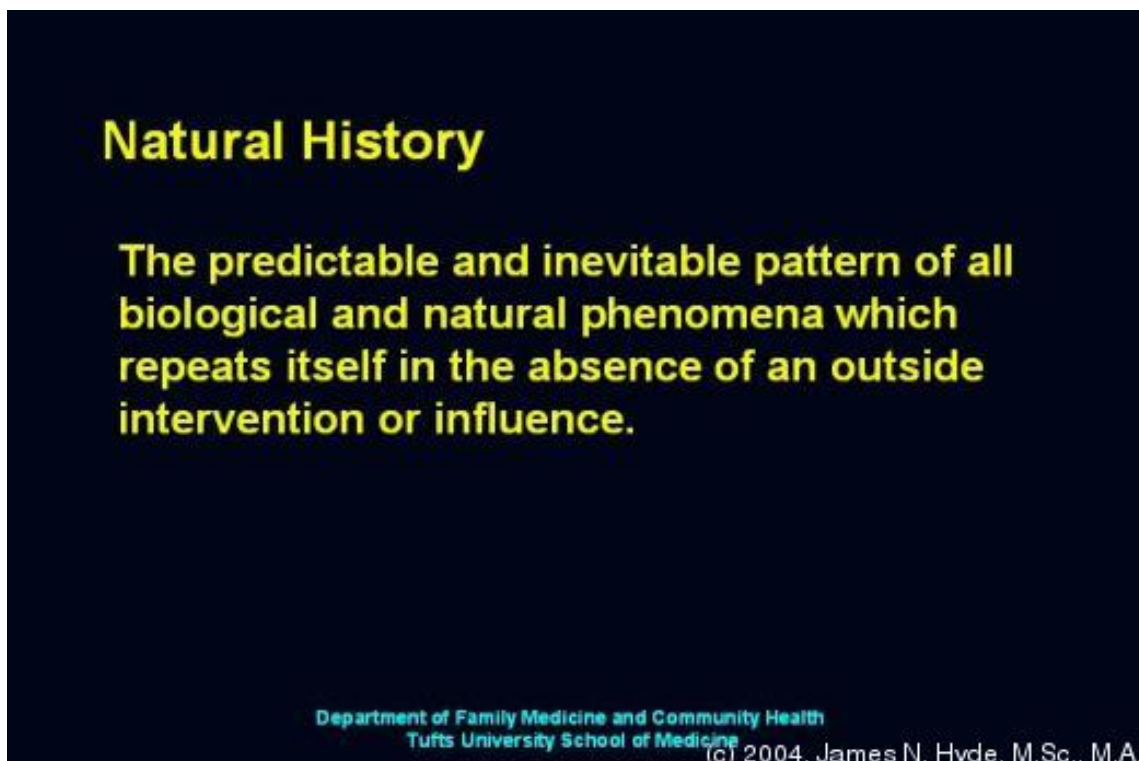


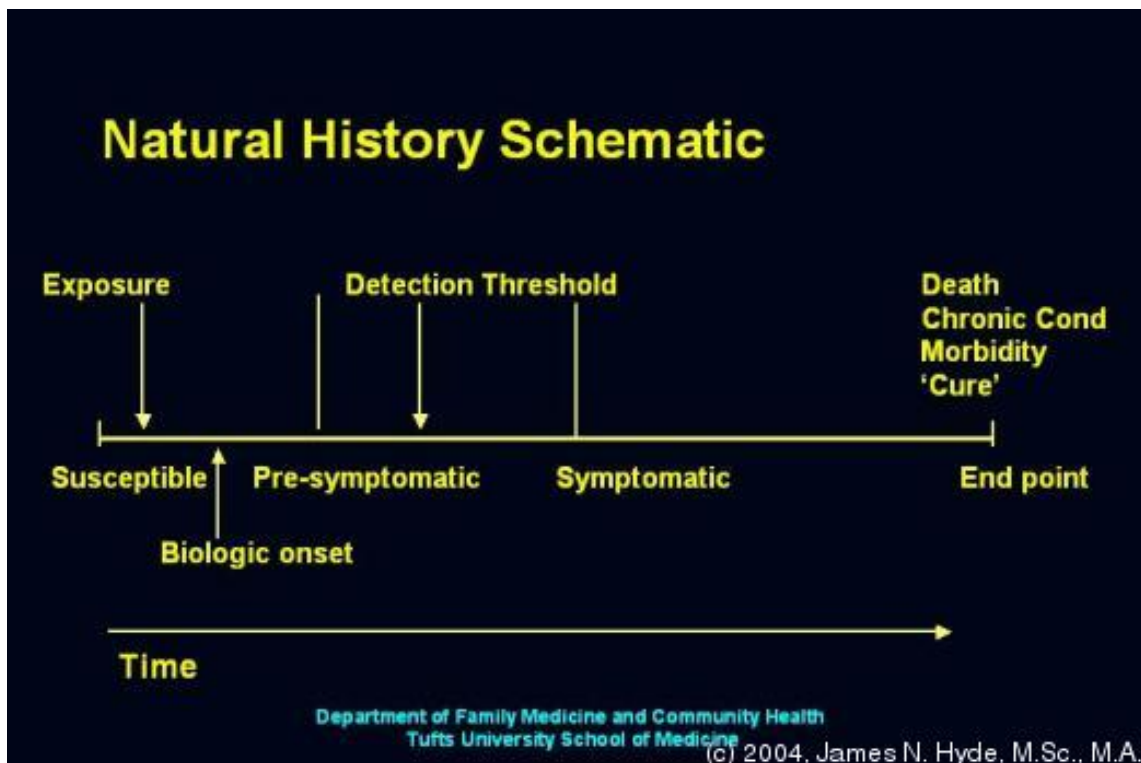
1. Lecture 3 - Introduction Slide



2. Natural History



3. Natural History Schematic



4. Natural History-Descriptors

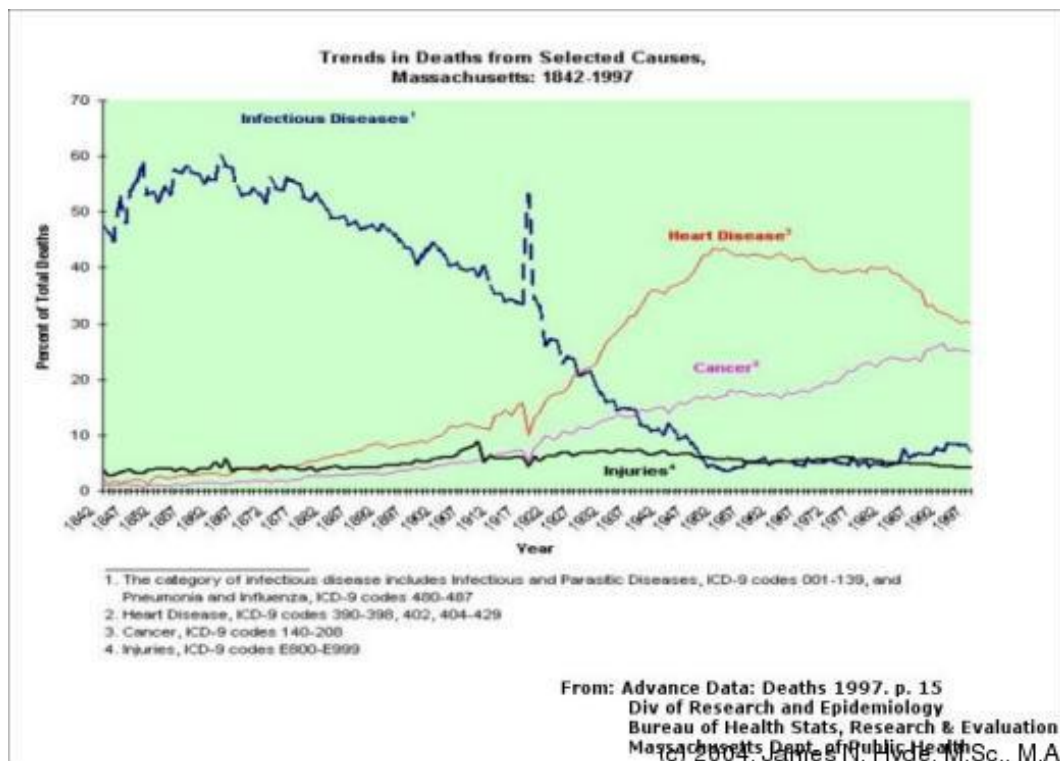
Natural History-Descriptors

- **Person** (age, sex, race, educational level, income family hx etc).
- **Place** (country, region, city town, location in a plant, etc).
- **Time** (time of year, day, progression over time)

Note: Understanding and mapping natural history gives clues to:
1) Etiology of dx, and 2) when and where to intervene.
(dx = diagnosis, hx = history)

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

5. Trends in Deaths from Selected Causes, MA: 1842-1887



6. Natural History-Descriptive Terms

Natural History-Descriptive Terms

- **Epidemic:** The occurrence in a community or region of a group of illnesses of a similar nature clearly in excess of normal expectancy. Epidemicity is therefore relative to usual frequency of disease in a region, among the same population, at the same time of year.
- **Endemic:** The constant presence of a disease within a given geographic area or population group.
- **Pandemic:** The constant presence of disease across a wide area or region.

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

7. Elements of Data Analysis

Descriptive Statistics
Elements of Data Analysis

- **Data Depiction**
- **Data Summarization/Reduction**
- **Data Transformation**

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

8. Rates

Descriptive Epidemiology
Rates

Rates can be thought of as a measure of the intensity with which events are occurring in a defined period of time.

$$\text{Rate} = \frac{\text{\# of events (e.g. deaths) during a defined time interval}}{\text{Population at risk during the same time interval}}$$

To calculate one needs:

- 1) a defined time interval (day, week, month year)
- 2) count of number of events occurring in the same interval)
- 3) estimate (or count) of population at risk in the same interval
- 4) a multiplier or constant (x 10; x 100; x 1000)

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

9. Rates-Example

Rates-Example

Between July 13 and July 21, 1995, 1177 deaths occurred in Chicago an 85% increase in the number that occurred in the same interval in 1994 (N=637). Of these deaths N=465 were identified as "heat related". The Census bureau estimated the Cook County population at 5,135,132 (1994) and 5,185,152 (1995).

One week mortality rates for Cook County (July 11-July 18) 1994 & 1995 respectively.

1994	$637/5,135,132 \times 100,000 = 12.4$	per 100,000 per week
1995	$1177/5,185,152 \times 100,000 = 21.2$	per 100,000 per week

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

10. Rates-Key Points

Rates-Key Points

- **Rates measure intensity with which events are occurring not magnitude**
- **Comparing rates is perilous unless groups are comparable**
- **The choice off the "constant" is determined by convention more than anything else**
- **Some commonly used rates have specific definitions**

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

11. Incidence- "cumulative incidence"

Incidence- "cumulative incidence"

$$\text{Incidence} = \frac{\text{\# of NEW cases of a disease in specified period}}{\text{\# of people at risk at the BEGINNING of the specified period}}$$

*Note: >Expressed as a proportion
>Specify time interval*

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

12. Cumulative Incidence-Examples

Cumulative Incidence-Examples

1) $\frac{22 \text{ NEW adolescent suicide attempts (NYC) during 1999}}{13,750 \text{ adolescents in NYC schools during 1999}}$
C.I. = .0016 = 0.16% per year

2) $\frac{1,450 \text{ NEW cases of flu South Florida Sept-Dec 1999}}{250,000 \text{ elderly living South Florida 1999}}$
C.I. = .0058 = 0.58% per 4 months

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

13. Incidence Density

Incidence Density

$$\text{Incidence Density} = \frac{\text{\# of NEW cases of a disease in specified period}}{\text{Sum of length of time at risk for each individual in the population}}$$

Note: Denominator expressed in "person-time" units e.g. person months, years etc.

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

14. Incidence Density: Example

Incidence Density: Example

Deaths post PCTA in 16 cases

Case #	Date of Procedure	Length F/U (mos)	Outcome (9/2000)
1	1/15/99	20	Alive
2	3/13/99	18	Deceased
3	9/13/99	12	Alive
·	·	·	·
·	·	·	·
·	·	·	·
15	3/13/00	8	Deceased
16	7/13/00	2	Alive
Total		240	

I.D. = 2 / 240 p-mos = .0083 or 8.3 per 1000 p-mos

Note: Study ended 9/13/2000

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

15. Incidence-Key Points

Incidence-Key Points

- Measure of disease onset-new cases
- Must specify time interval
- Only population at risk (theoretically) is in the denominator-practically it can be a problem
- Incidence density provides a more precise estimate of risk-not always possible to calculate

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

16. Prevalence-Example

Prevalence-Example

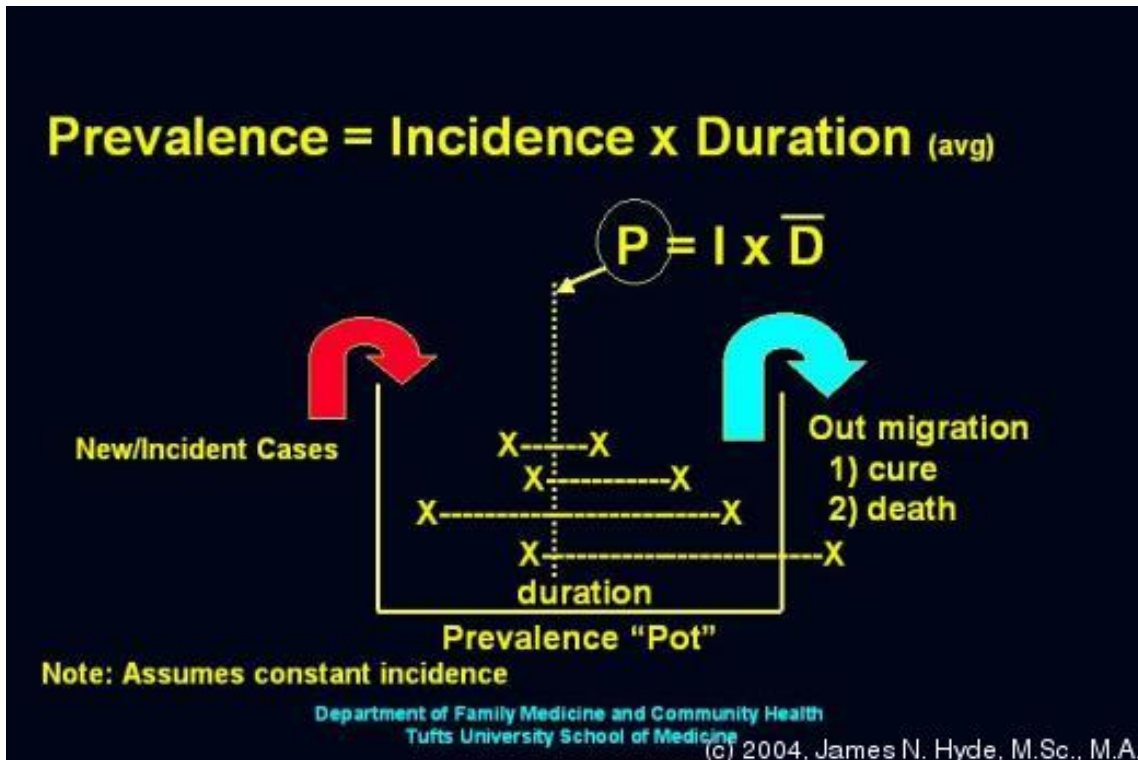
10 first year medical students have URIs

165 first year medical students

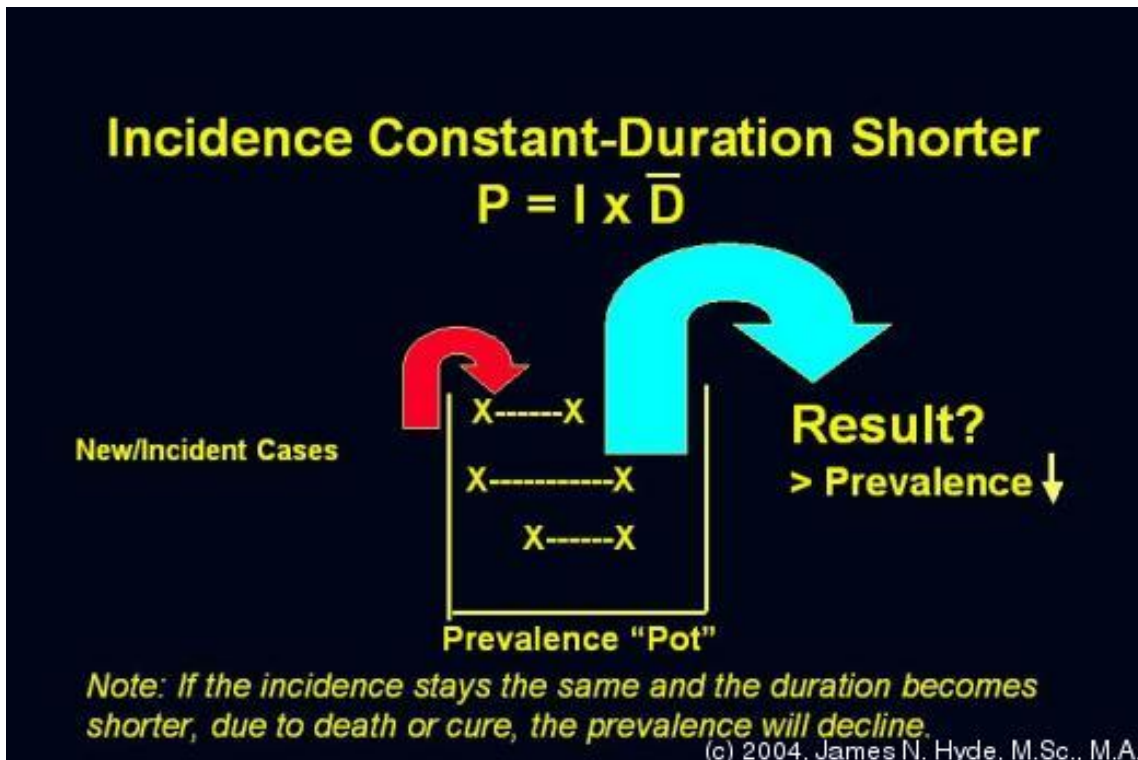
$P = 10/165 = .0606 = 6.06\% \text{ or } 6.06 \text{ per } 100$

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

17. Prevalence = Incidence x Duration (avg)



18. Incidence Constant-Duration Shorter



19. Incidence Increases/Duration Constant

Incidence Increases/Duration Constant

**Note: If the incidence increases and duration stays the same
The prevalence will increase.**

(c) 2004, James N. Hyde, M.Sc., M.A.

20. Incidence of Type II Diabetes Two Towns 1999

Using Rates to make Comparisons Incidence of Type II Diabetes Two Towns 1999

Age Group	Town 'A'			Town 'B'		
	Pop	# New Cases	Rate	Pop	# New Cases	Rate
<35	2000	24	12/1000	5000	60	12/1000
35-64	3000	72	24/1000	3000	72	24/1000
≥ 65	5000	180	36/1000	2000	72	36/1000
TOTAL	10,000	276	27.6/1000	10,000	204	20.4/1000

Rate ratio = 27.6/20.4 = 1.35

Department of Family Medicine and Community Health
Tufts University School of Medicine

(c) 2004, James N. Hyde, M.Sc., M.A.

21. Rate Adjustment (direct method)

Rate Adjustment (direct method)

Standard Population	Rate for A	Expected # A	Rate for B	Expected # B
7000	12/1000	84	12/1000	84
6000	24/1000	144	24/1000	144
7000	36/1000	252	36/1000	252
20,000	24/1000	480	24/1000	480

Adjusted rate ratio = 24 per 1000/24 per 1000 = 1

(c) 2004, James N. Hyde, M.Sc., M.A.

22. Types of Rates

Types of Rates

Age Group	Town 'A'			Town 'B'		
	Pop	# New Cases	Rate	Pop	# New Cases	Rate
<35	2000	24	12/1000	5000	60	12/1000
35-64	3000	72	24/1000	3000	72	24/1000
≥ 65	5000	180	36/1000	2000	72	36/1000
Total	10,000	276	27.6/1000	10,000	204	20.4/1000

↙

category specific rate

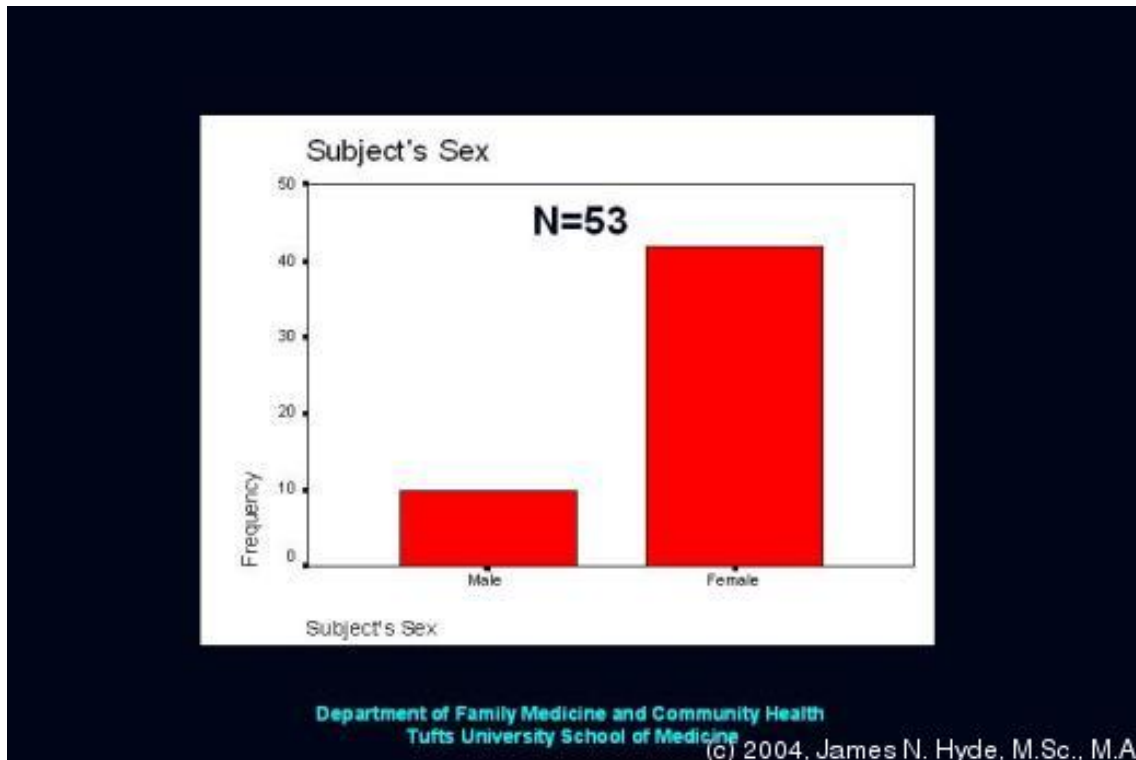
↘

crude rate

Rate Ratio = 27.6/20.4 = 1.35

(c) 2004, James N. Hyde, M.Sc., M.A.

23. Frequency/Gender



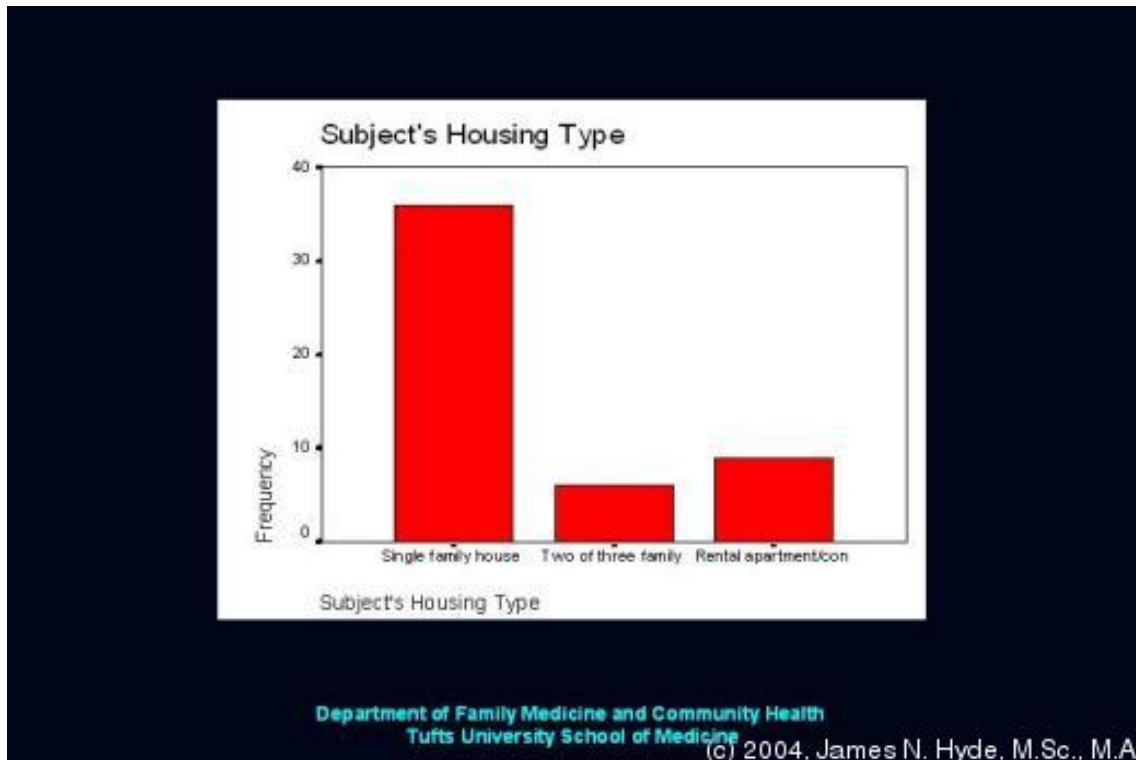
24. Osteoarthritis of Knees and Obesity

Osteoarthritis of Knees and Obesity

Researchers wanting to study the relationship between obesity and osteoarthritis (OA) of the knees, recruited 20000 women to participate in a 15 year follow-up study. Subjects were screened at the beginning of the study for any evidence of existing OA of the knees as well as for obesity. Those who showed any existing level of OA were eliminated from the study. The subjects were subsequently examined annually for evidence of OA and classified as either OA + or OA -. Their heights and weights were also measured. At the conclusion of the study 240 of the 8,000 women who were classified as obese had been diagnosed as having OA , while 130 of the normal weight women showed evidence of OA of the knees.

Department of Family Medicine and Community Health
Tufts University School of Medicine
(C) 2004, James N. Hyde, M.Sc., M.A.

25. Frequency/Housing Type



26. Contingency Tables 2x2 Table

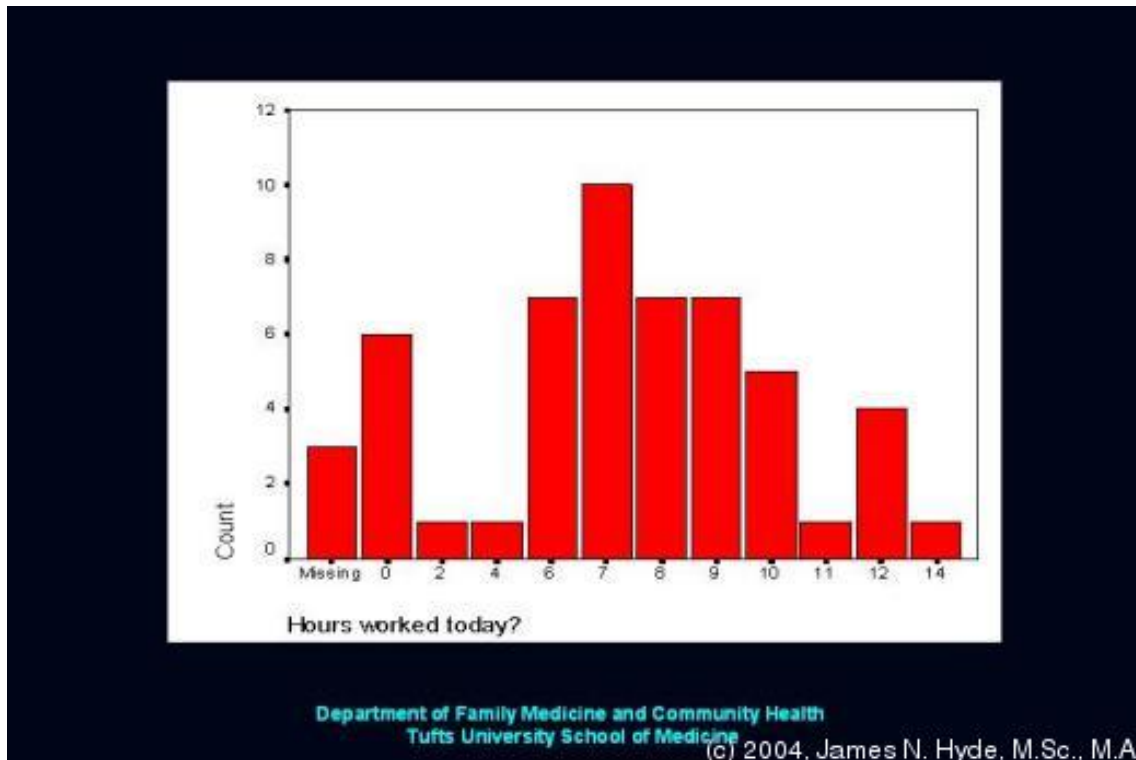
Contingency Tables 2x2 Table

Outcome

Exposure	Y(+)	N(-)	Totals
Y(+)	a	b	a+b
N(-)	c	d	c+d
Totals	a+c	b+d	N=a+b+c+d

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

27. Count/Hours Worked



28. Osteoarthritis of the knees and Obesity N=20,000 women

**Osteoarthritis of the knees and Obesity
N=20,000 women**

Exposure	Y(+)	N(-)	Totals
Y(+)	240	7,760	8,000
N(-)	130	11,870	12,000
Totals	370	19,630	20,000

Rate ratio = $\frac{240/8,000}{130/12,000} = .03/.01 = 3$

Department of Family Medicine and Community Health
Tufts University School of Medicine
(C) 2004, James N. Hyde, M.Sc., M.A.

29. Subject's Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-21 yrs	11	20.4	20.8	20.8
	22-30 yrs	8	14.8	15.1	35.8
	31-40 yrs	17	31.5	32.1	67.9
	41-50 yrs	14	25.9	26.4	94.3
	51-60 yrs	3	5.6	5.7	100.0
	Total		53	98.1	100.0
Missing	9	1	1.9		
Total		54	100.0		

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

30. Relative Risk

Relative Risk

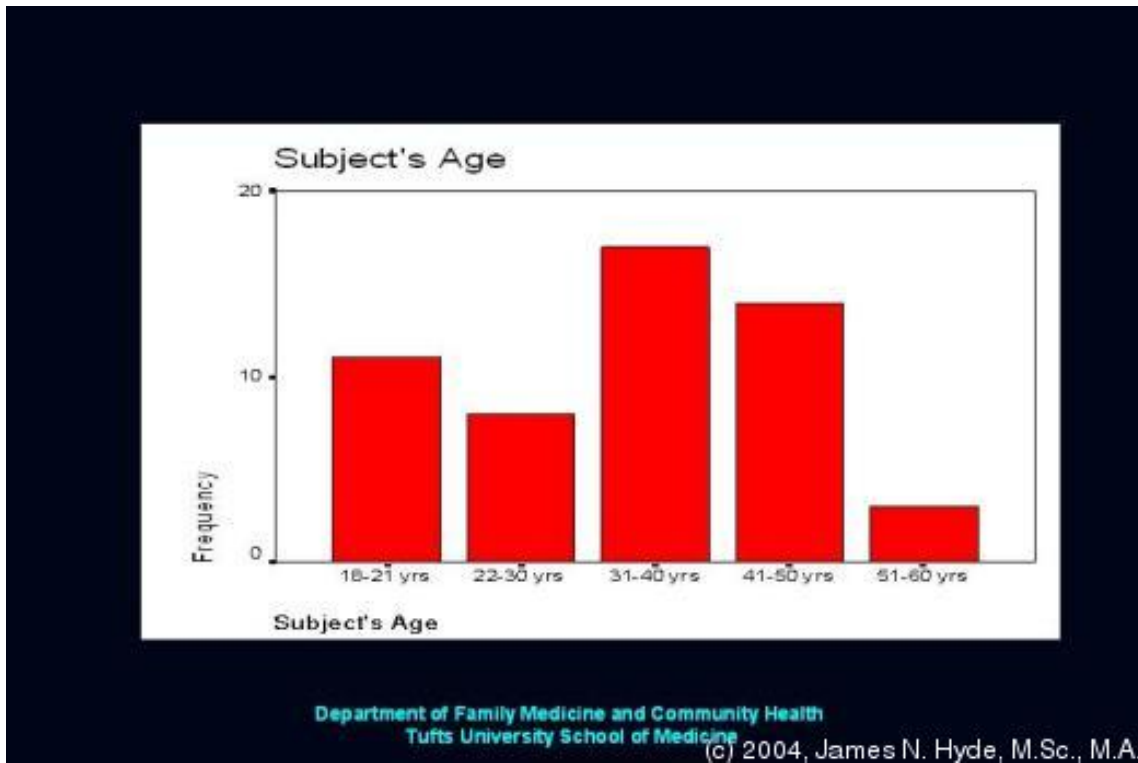
$$\frac{\text{Incidence in exposed}}{\text{Incidence in non-exposed}}$$

Key Points:

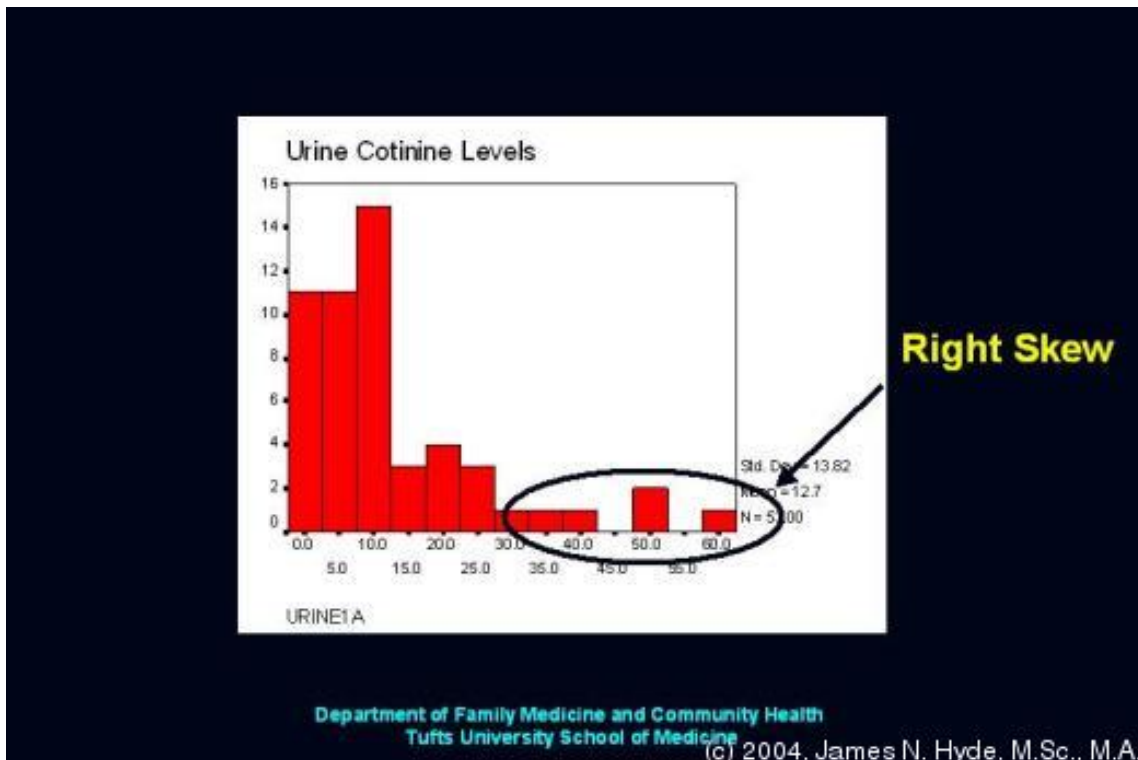
- 1) measure of strength of association
- 2) can be greater or less than 1
- 3) dimensionless

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

31. Subject's Age/Frequency



32. Urine Cotinine Levels



33. Odds Ratio: (in Case/Control Studies)

Odds Ratio: (in Case/Control Studies)

Case Control

Exposure	(+)	(-)
Y(+)	a	b
N(-)	c	d
Totals	a + c	b + d

OR = a x d / b x c

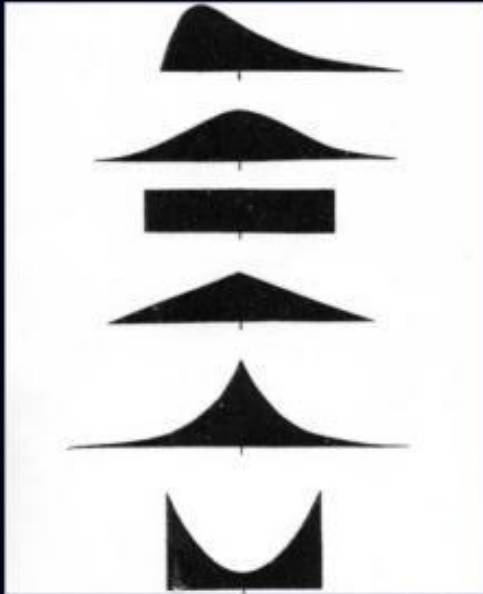
Note:

- 1) OR is interpreted the same way as Relative Risk
- 2) OR are also seen in other types of studies as the result of a type of statistical analysis called logistic regression

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

34. Selection of Types of Distributions

Selection of Types of Distributions



Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

35. Attributable Risk

Attributable Risk

$$AR = \text{Inc in Exposed (+)} - \text{Inc in Exposed (-)}$$

From OA Example:

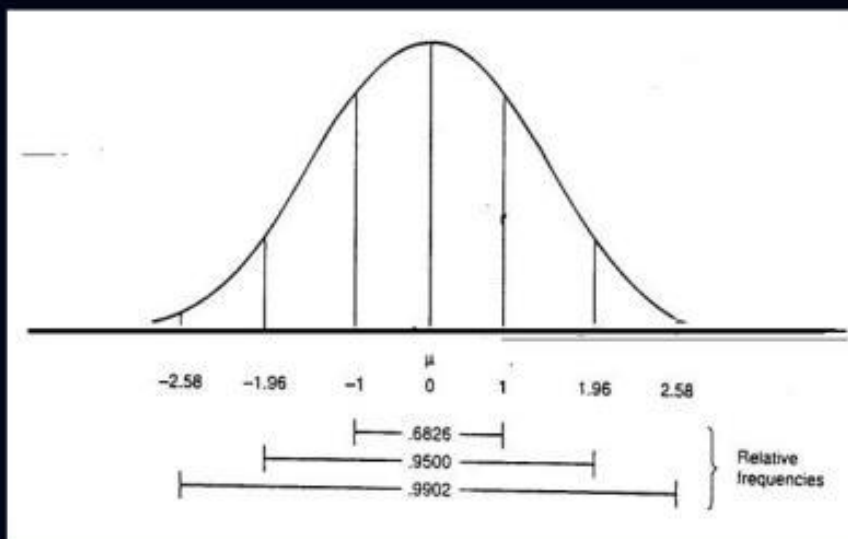
.03-.01 or .02 (or 20 per 1,000 per 15 years)

Note: AR is a measure of disease burden

Department of Family Medicine and Community Health
Tufts University School of Medicine

(C) 2004, James N. Hyde, M.Sc., M.A.

36. Key Point

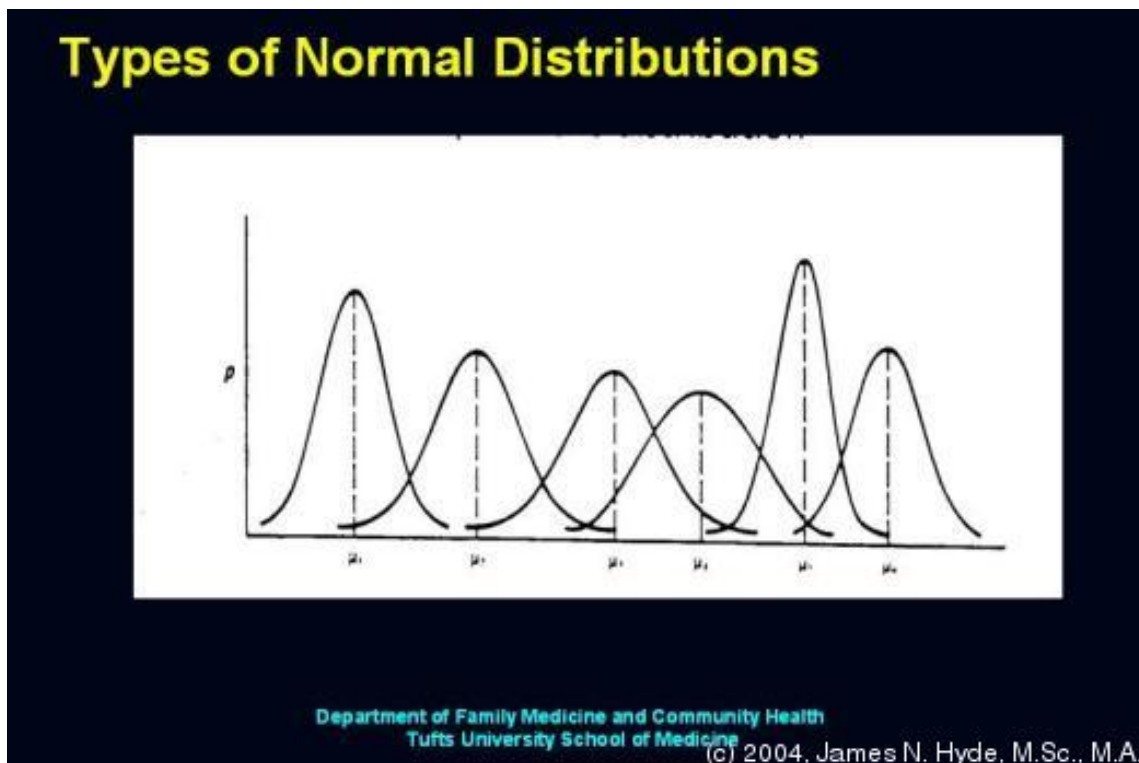


Key Point: the 68-95-99% rule

Department of Family Medicine and Community Health
Tufts University School of Medicine

(C) 2004, James N. Hyde, M.Sc., M.A.

37. Types of Normal Distributions



38. Summary: Descriptive Epidemiology

- ### Summary: Descriptive Epidemiology
- Descriptive studies as first step in unraveling etiology
 - Natural history clues to etiology and prevention
 - Rates (category specific, crude, adjusted)
 - Adjustment essential for comparing rates
 - Key epi measures: Incidence (cumulative and incidence density) and prevalence
 - Relationship between incidence, prevalence and duration
 - Relative risk (odds ratio) and attributable risk
- Department of Family Medicine and Community Health
Tufts University School of Medicine
(C) 2004, James N. Hyde, M.Sc., M.A.

39. Prevalence

Prevalence

$$\text{Prevalence} = \frac{\text{\# of cases with disease at a specified period}}{\text{\# of people in the population at specified time}}$$

Note: 1) includes OLD and NEW cases
2) prevalence is a proportion i.e. percent, fraction
3) cross-sectional measure
4) measure of disease burden

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

40. Notation used for Populations vs Samples

Notation used for Populations vs Samples

Characteristic	Population	Sample
Proportion	π	P
Mean	μ	\bar{X}
Variance	σ^2	s^2
Standard Deviation	σ	s

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

41. Natural History-Key Points

Natural History-Key Points

- There are discernible patterns to health events
- What happens in the absence of intervention
- Understanding patterns provides clues to etiology and prevention
- Not predictive of individual outcome

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

42. Central Tendency

Mean Central Tendency

$$\bar{X} = \frac{\sum x_i}{N}$$

Data set: 18, 21, 17, 13, 15 Mean= 16.8

Median

1. Order data smallest to largest values
2. For odd number take middle value
3. For even number take mean of two middle values

Data set: 18, 21, 17, 13, 15 Median= 17

Key Points:

- 1) When data are symmetric mean=median=mode
- 2) When data are NOT symmetric mean≠median≠mode

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

43. Spread

Spread

Variance

$$s^2 = \frac{\sum (\bar{X} - x_i)^2}{(N-1)}$$

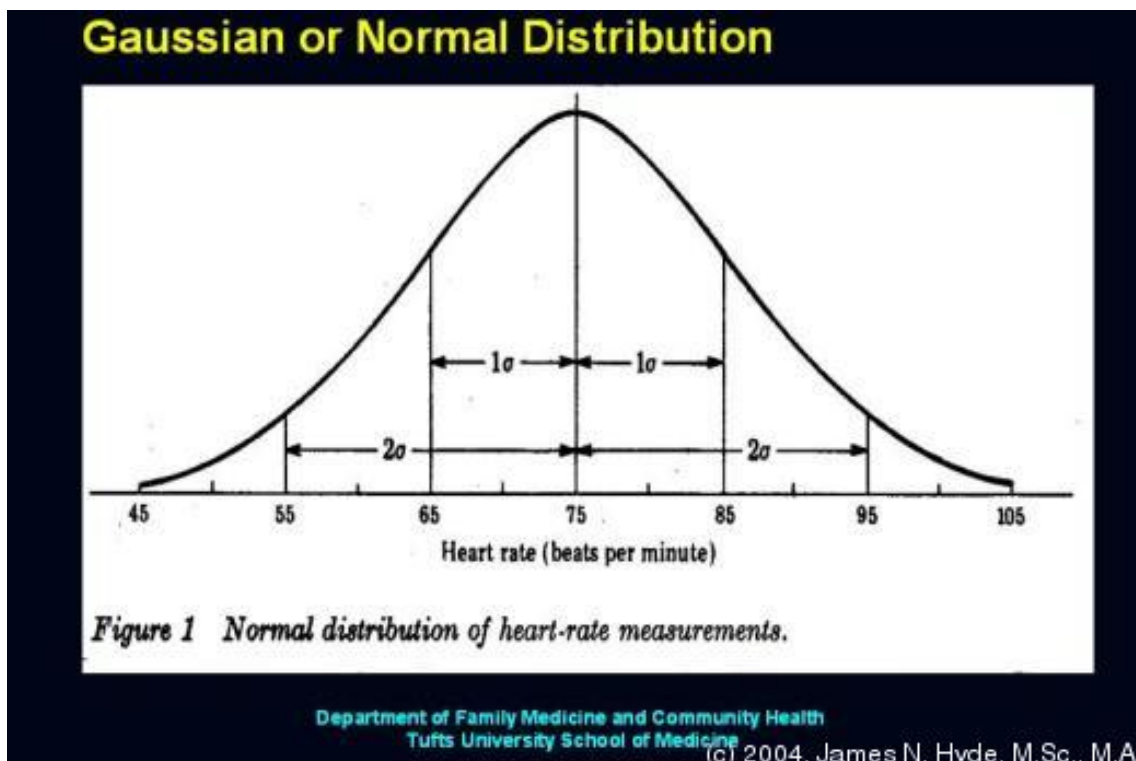
Standard Deviation

$$s = \sqrt{\frac{\sum (\bar{X} - x_i)^2}{(N-1)}}$$

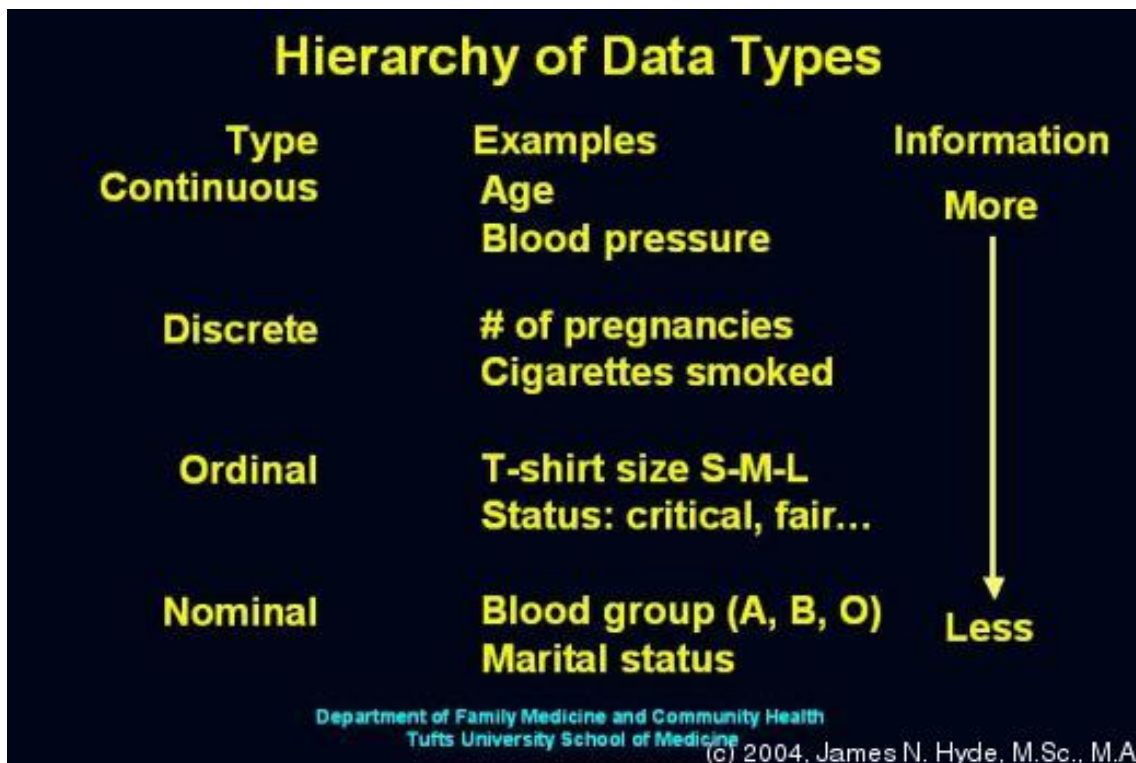
Data set: 18, 21, 17, 13, 15
Range = 7
 $s^2=9.19$ $s=3.01$

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

44. Gaussian or Normal Distribution



45. Hierarchy of Data Types



46. z scores

Z scores

$$z = \frac{x_i - \mu}{\sigma}$$

Where: x_i = raw score
 μ = population mean
 σ = standard deviation
z = 'z score' the 'raw' score transformed to the number of standard deviation units the score is from the mean

z is a measure of the distance in sd units of a value x_i from the mean.

47. Concept of risk

Concept of risk

The probability or likelihood of moving from one health "state" to another i.e. the risk of dying or contracting West Nile virus.

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.

48. z score example

z score example

If heart rate = 85 BPM, what percentile does that represent?

$$z = \frac{85 - 75}{10}$$
$$z = 1.0$$
84 percentile**

Figure 1 Normal distribution of heart-rate measurements.

Note: we have "transformed" data from one measurement System to another i.e. BPM to percentile.

** values can be found in z table (standard normal distribution) Pagano.

Department of Family Medicine and Community Health
Tufts University School of Medicine
(c) 2004, James N. Hyde, M.Sc., M.A.