Design IV

E232 Fall 07

Class 15

Bruce McNair
bmcnair@stevens.edu

Today's topics

- Statistical Analysis of Experimental Data

Representative Experimental Data

- Student grades in E232, Fall 2007 on Quiz 1:

Grade	Number of students with this grade(N)
121.5	1
118.75	2
118.125	1
117.5	1
116.2	2
115	1
113.75	1
112.5	2
111.25	3
110	1
107.5	1
106.25	1
105	3
103.75	1
102.5	3

Grade	N
101.25	4
100	1
98.75	1
97.5	4
96.25	4
95	3
92.5	2
91.875	1
90	4
88.75	3
86.875	3
85	4
83.75	5
82.5	2
81.25	1

Grade	N
80	3
78.75	2
77.5	3
76.25	1
75	2
73.75	1
72.5	2
71.25	2
70	1
62.5	1
60	2
57.5	1
56.875	1
55	1

Representative Experimental Data

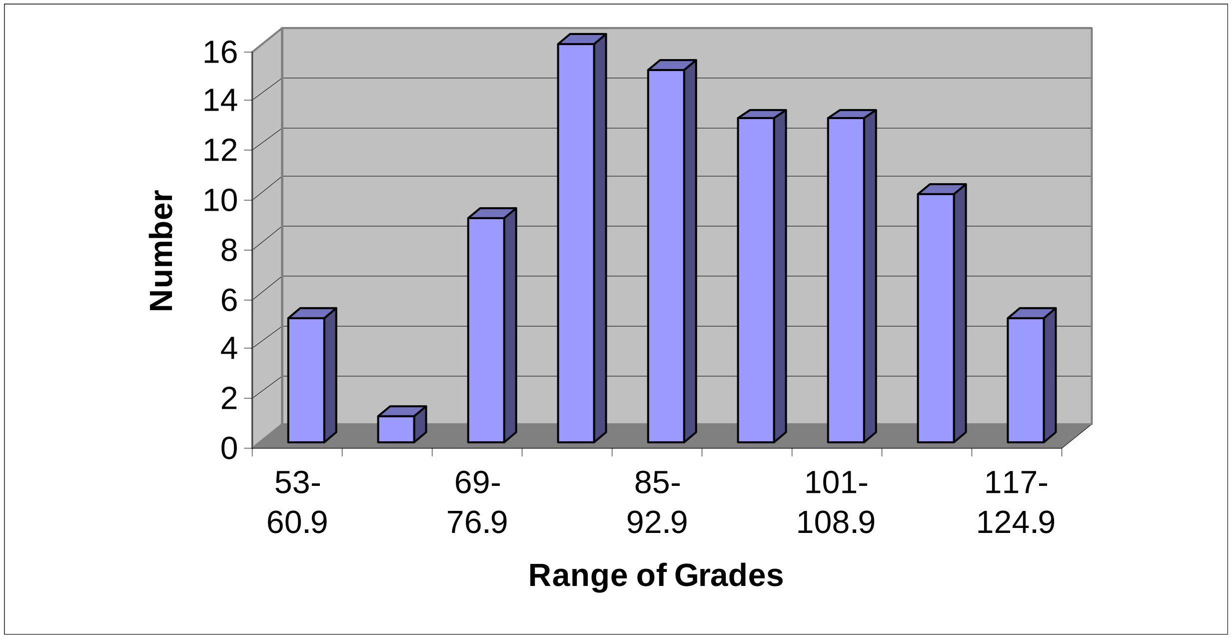
- Student grades in E232, Fall 2007 on Quiz 1:

Grade range	Number of students in this range of grades
117-124.9	5
109-116.9	10
101-108.9	13
85-92.9	15
77-84.9	16
69-76.9	9
61-68.9	1
53-60.9	5

- **Experimental details are lost with “bins” but trends are easier to see.**
- **As a rule of thumb, the number of “bins” should be approximately the square root of the number of observations (students in this case)**

Representative Experimental Data

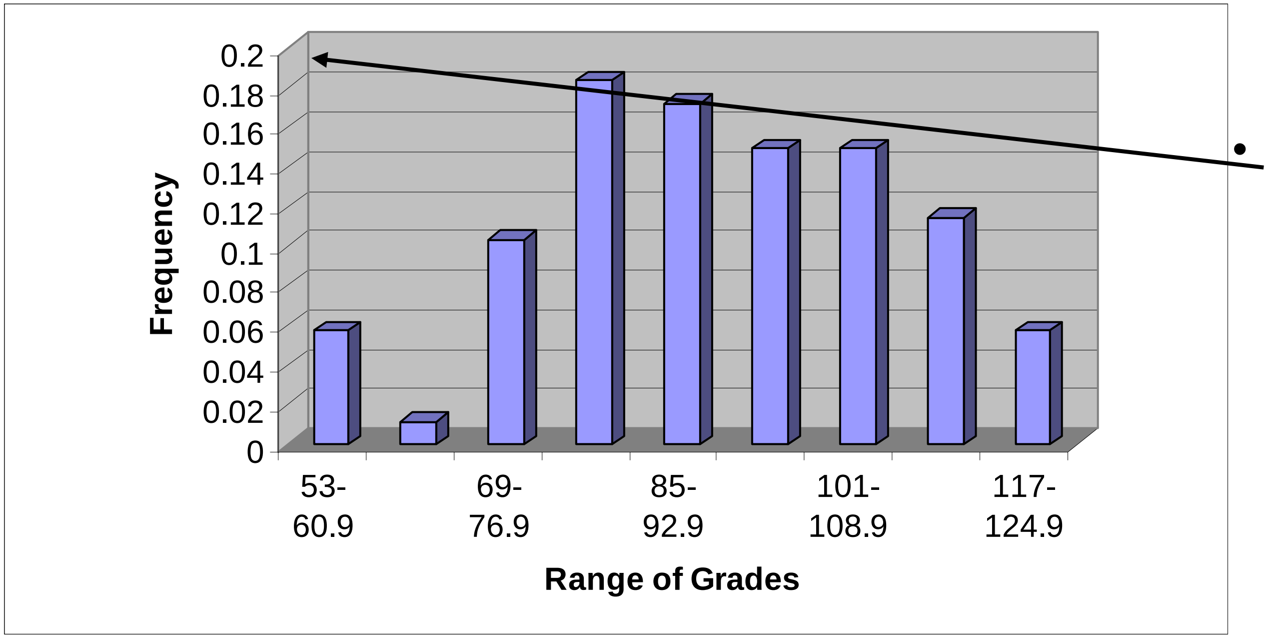
- Student grades in E232, Fall 2007 on Quiz 1:



- **Trends are much easier to spot with graphical (histogram) representation**

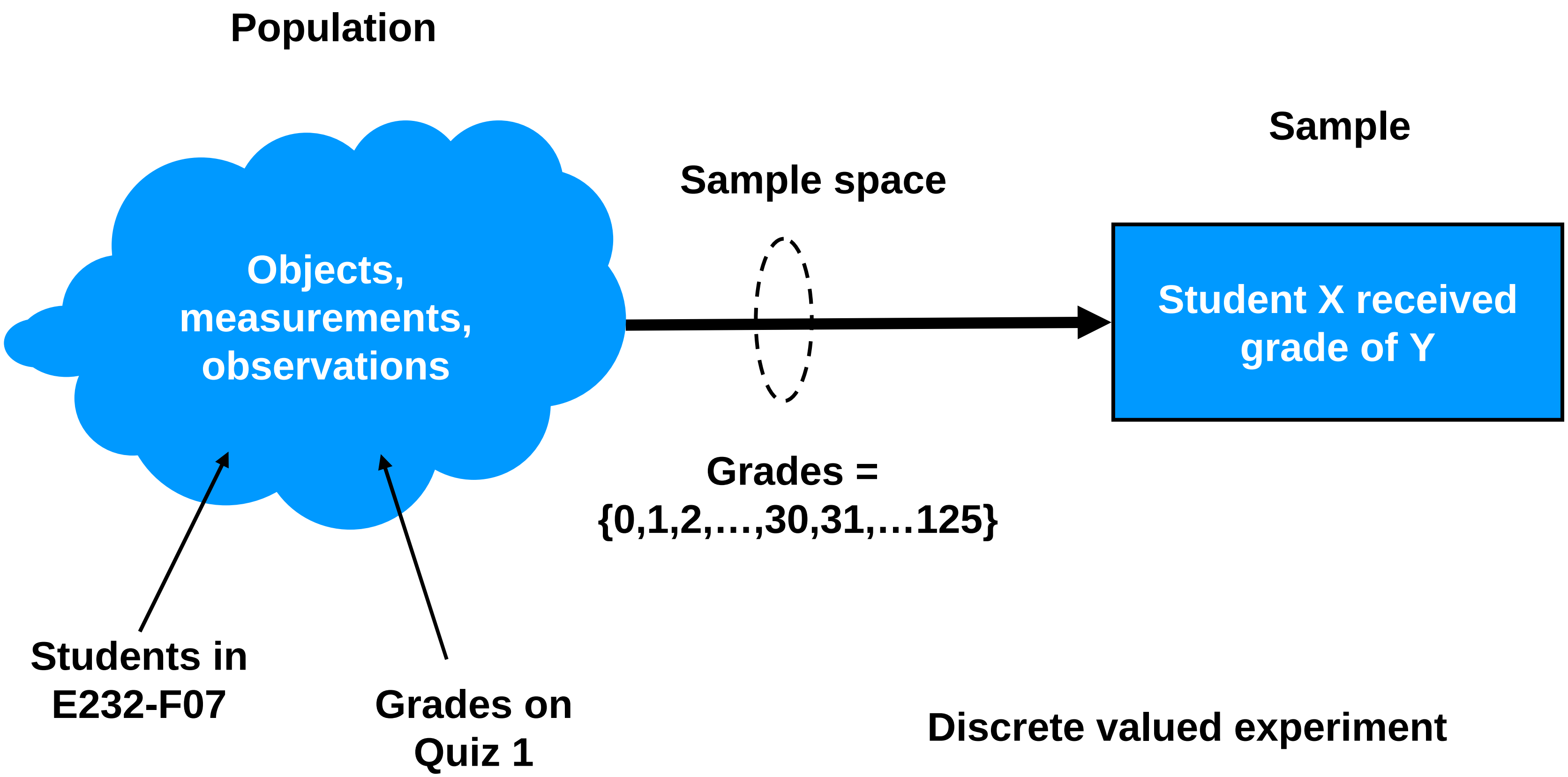
Representative Experimental Data

- Student grades in E232, Fall 2007 on Quiz 1:

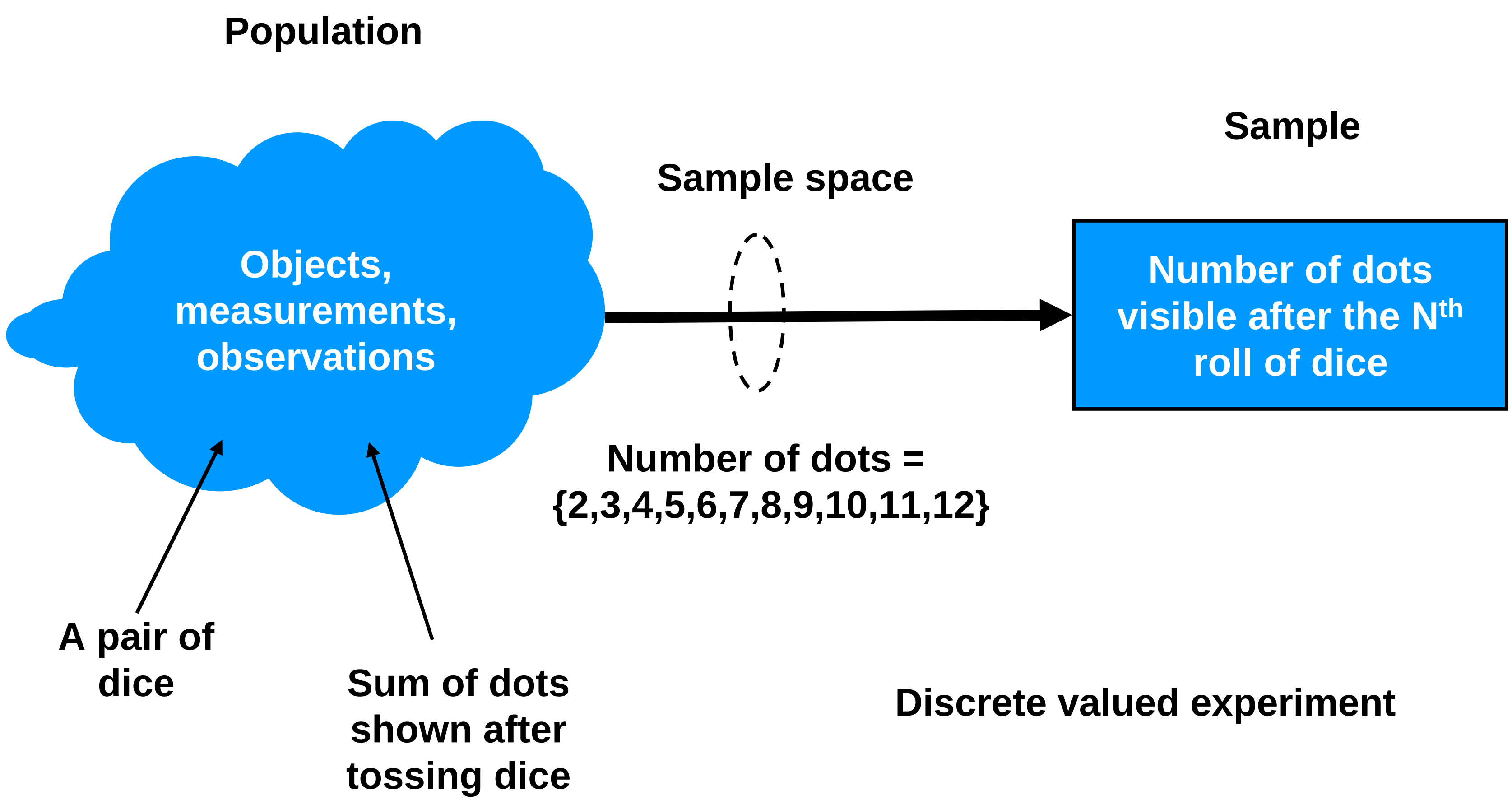


- **By normalizing the graph to the number of observations, the plot becomes generic for all experimental data set sizes**

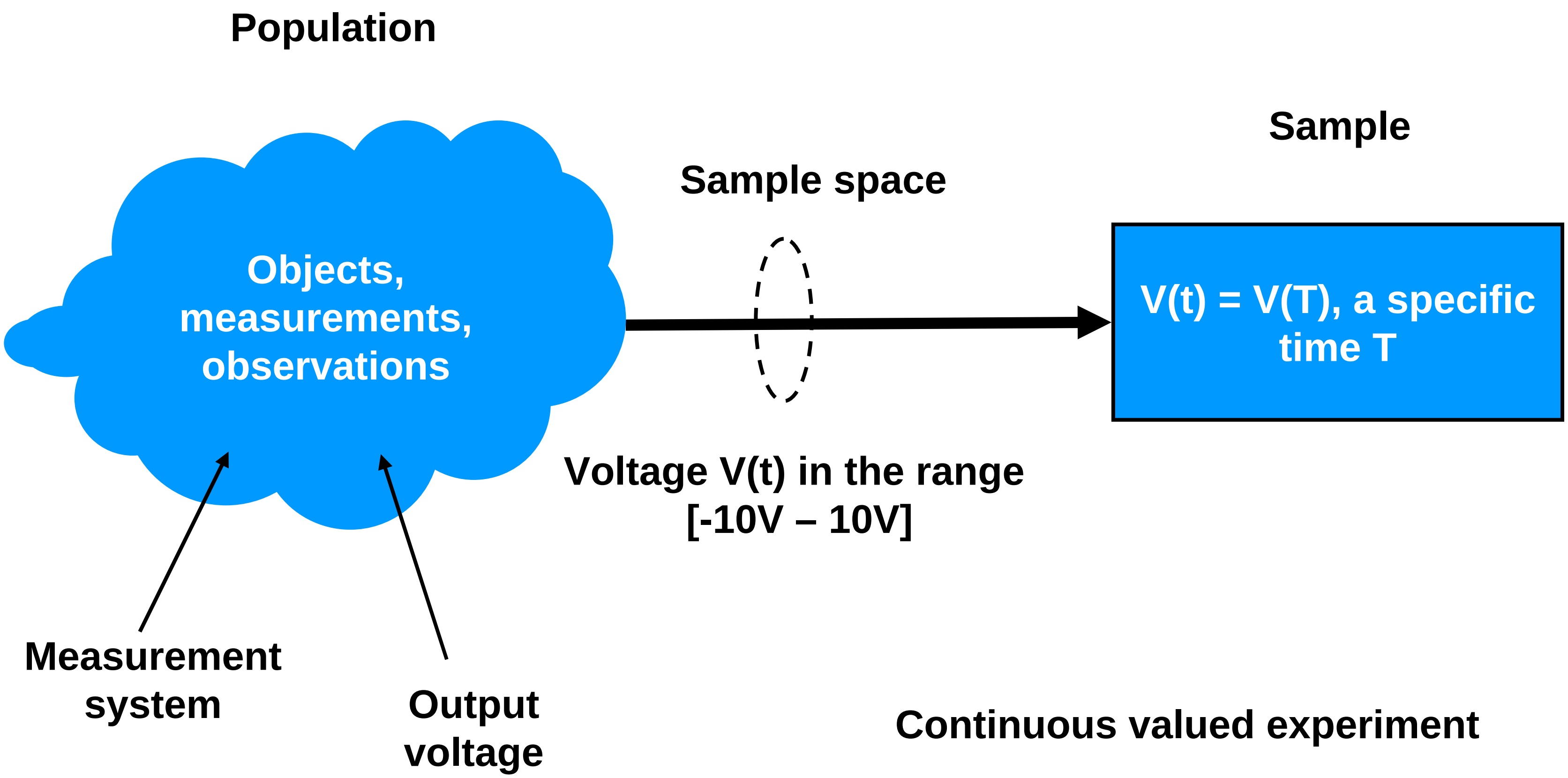
Terminology



Terminology

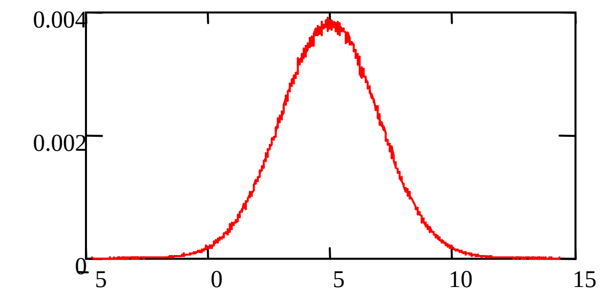


Terminology

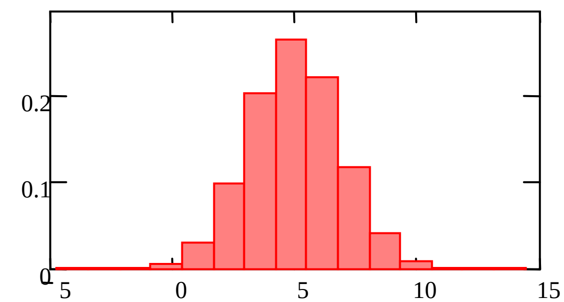


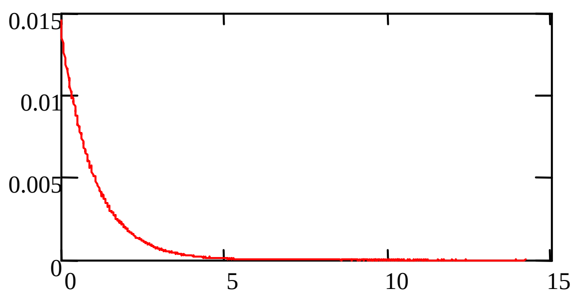
Continuous valued experiment

Sample Distributions

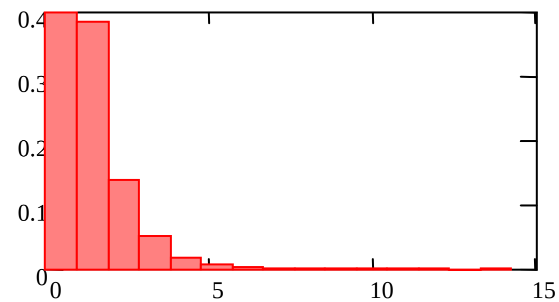


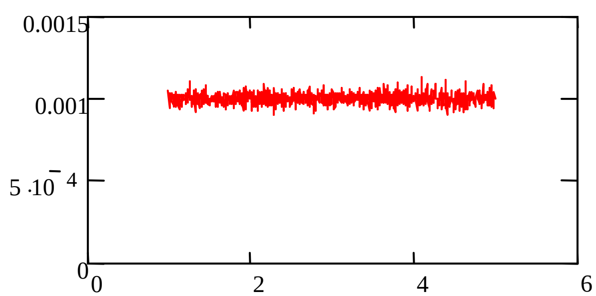
Normal/ Gaussian



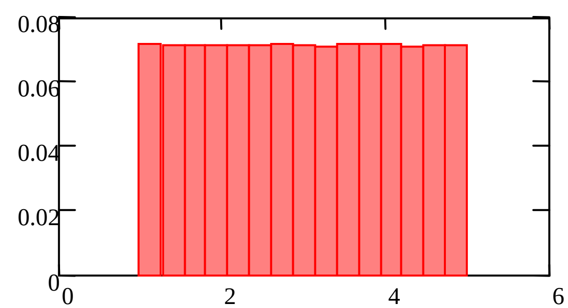


Exponential





Uniform

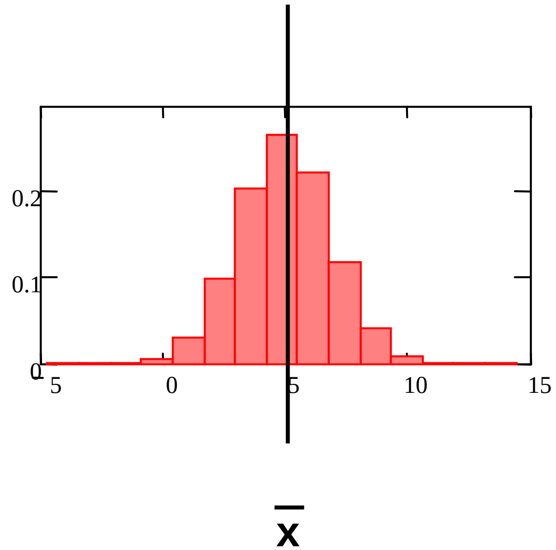


Continuous **Discrete**

Characterizing Distributions

- Mean, average, central tendency

- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**



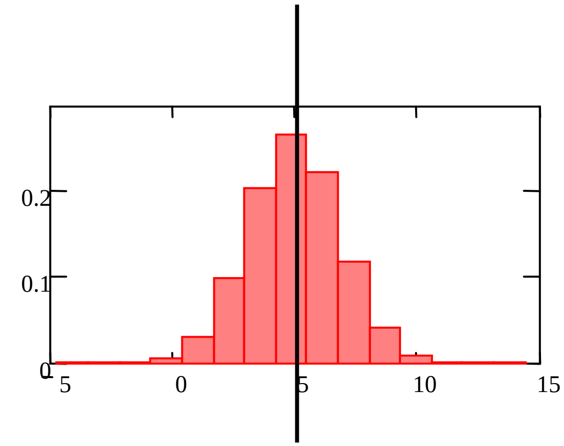
$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} = \sum_{i=1}^{n} \frac{x_i}{n}$$

Average of a set of samples

Characterizing Distributions

- Mean, average, central tendency

- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**



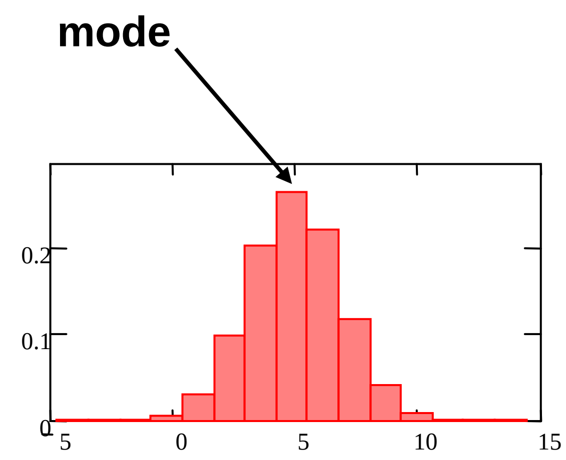
μ

$$\mu = \frac{x_1 + x_2 + x_3 + \ldots + x_N}{N} = \sum_{i=1}^{N} \frac{x_i}{N}$$

Mean of a population

Characterizing Distributions

- Mode – most frequent value

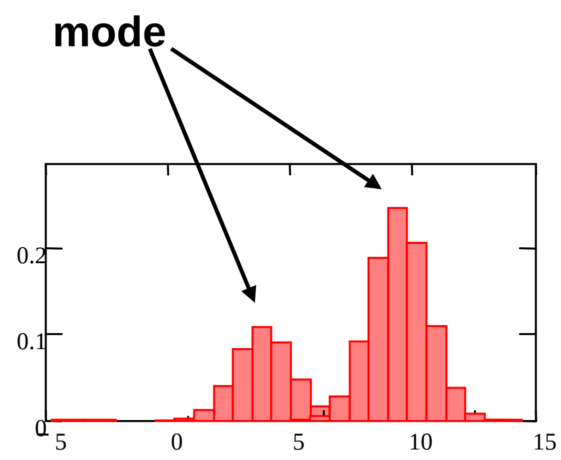


- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**

Characterizing Distributions

- Mode – most frequent value

- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**



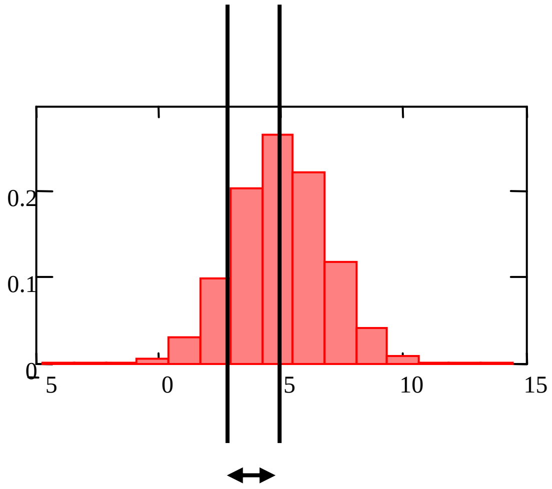
Bimodal distribution

Characterizing Distributions

- Measuring deviation

- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**

$$d_i = x_i - \overline{x}$$



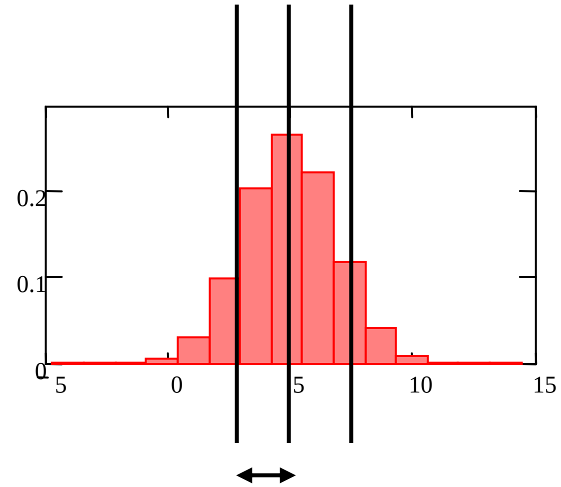
Deviation of a measurement from average

Characterizing Distributions

- Measuring deviation

- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**

$$\sigma = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \mu)^2}{N}}$$



Population standard deviation
σ

Standard deviation from mean of population (note resemblance to RMS)

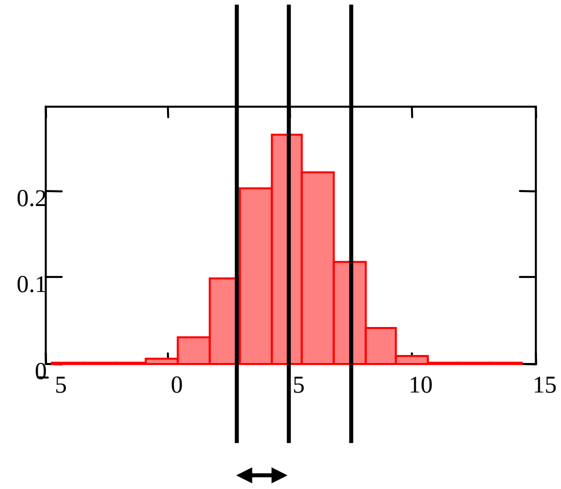
Characterizing Distributions

- Measuring deviation

- **Assume a set of n measurements**
- **Assume a population of N elements**
- **Assume measurement values, x_i**

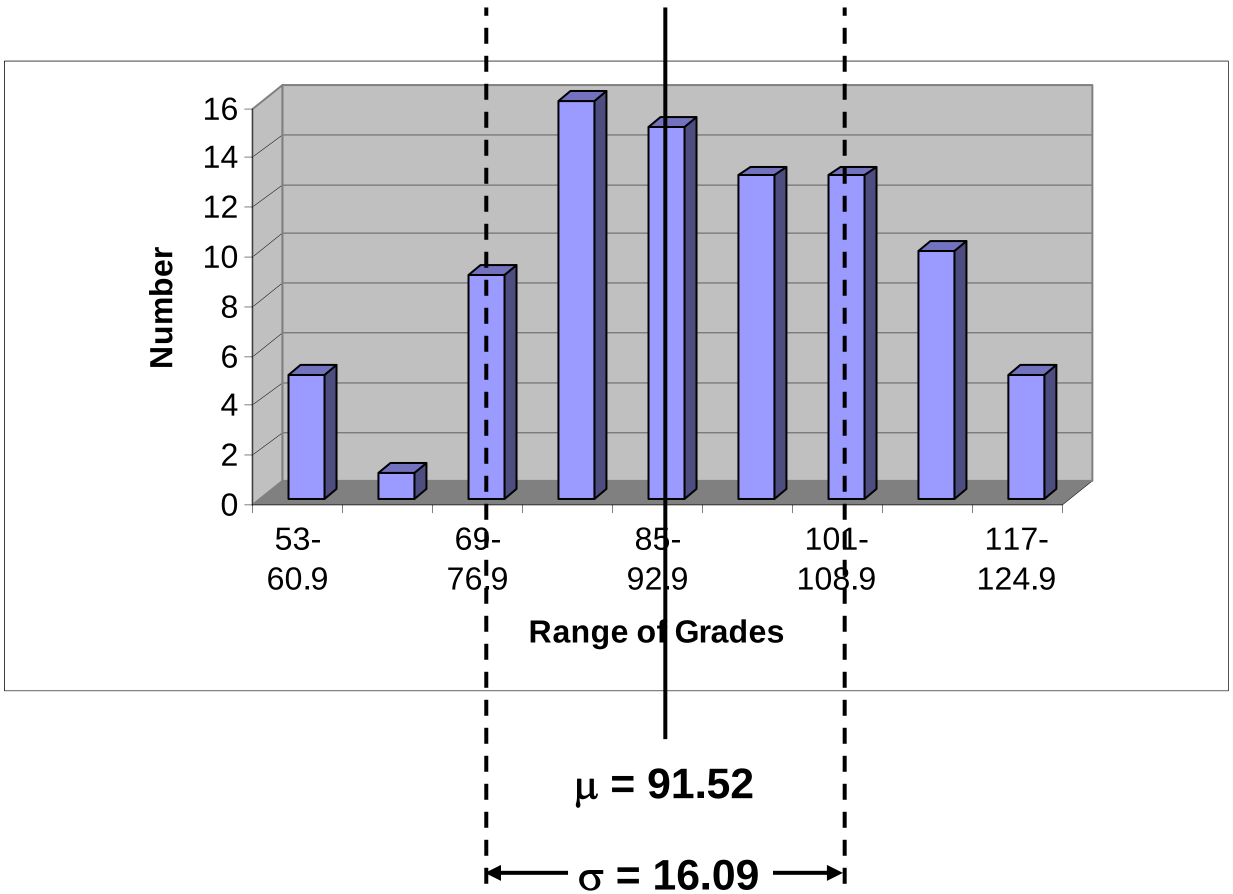
$$S = \sqrt{\sum_{i=1}^{n} \frac{\left(x_i - \overline{x}\right)^2}{(n-1)}}$$

Sample standard deviation when mean is not known in advance

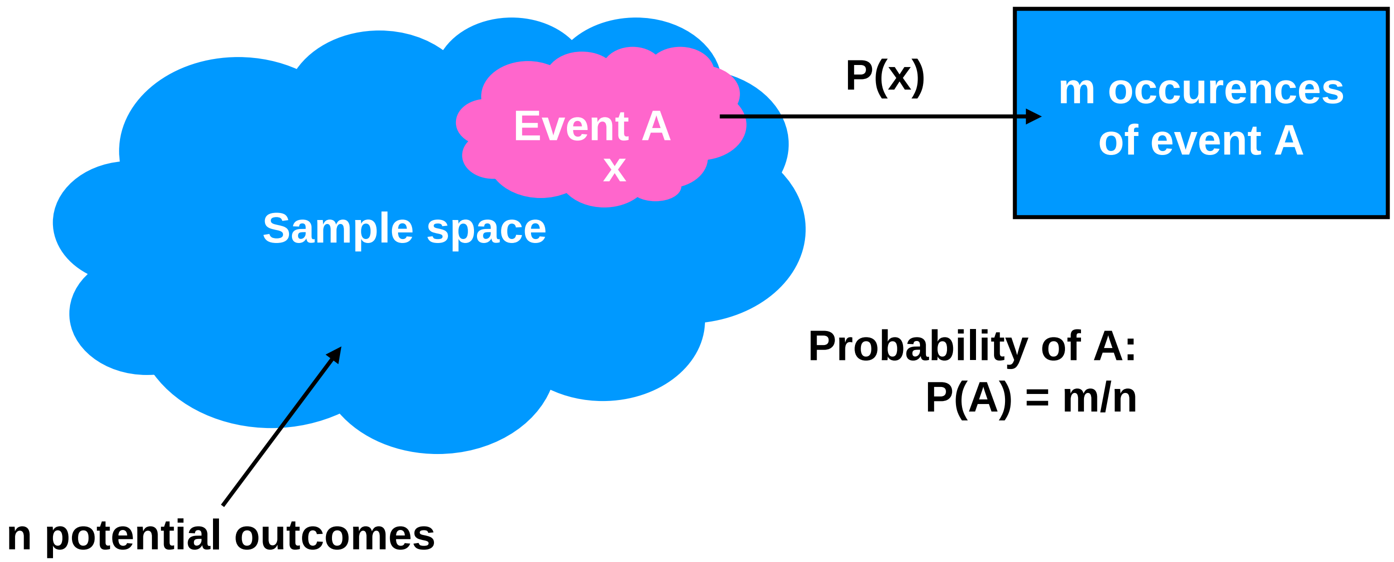


Sample standard deviation S

Statistics For Homework Class Grades

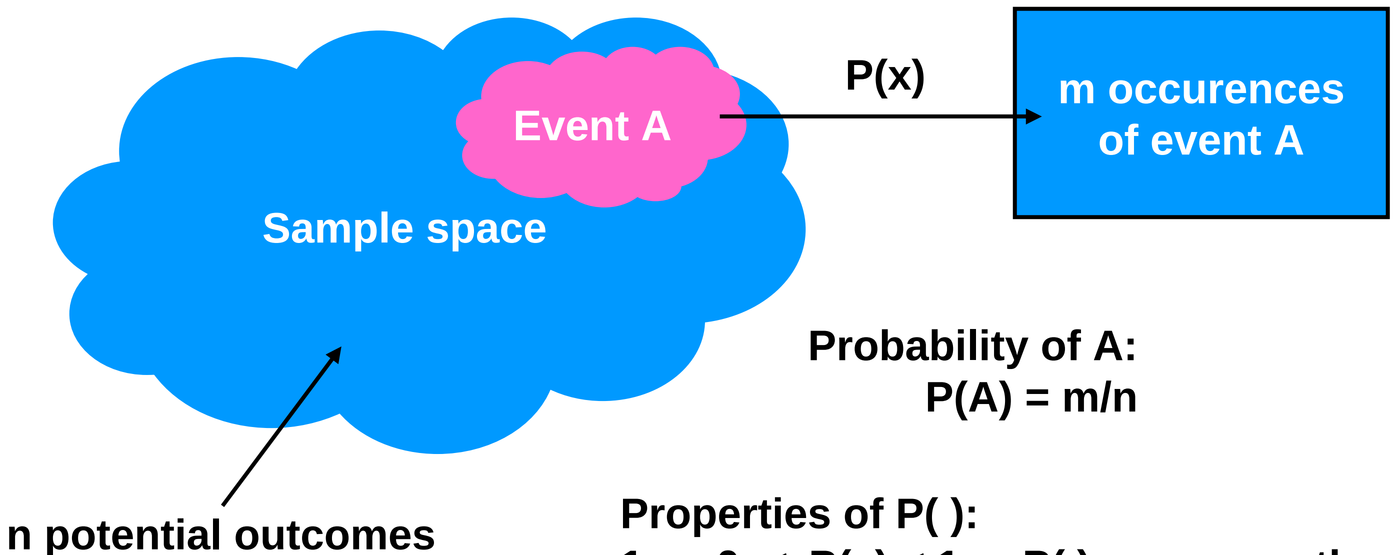


Probability Measures



Probability of A:
P(A) = m/n

Probability Measures



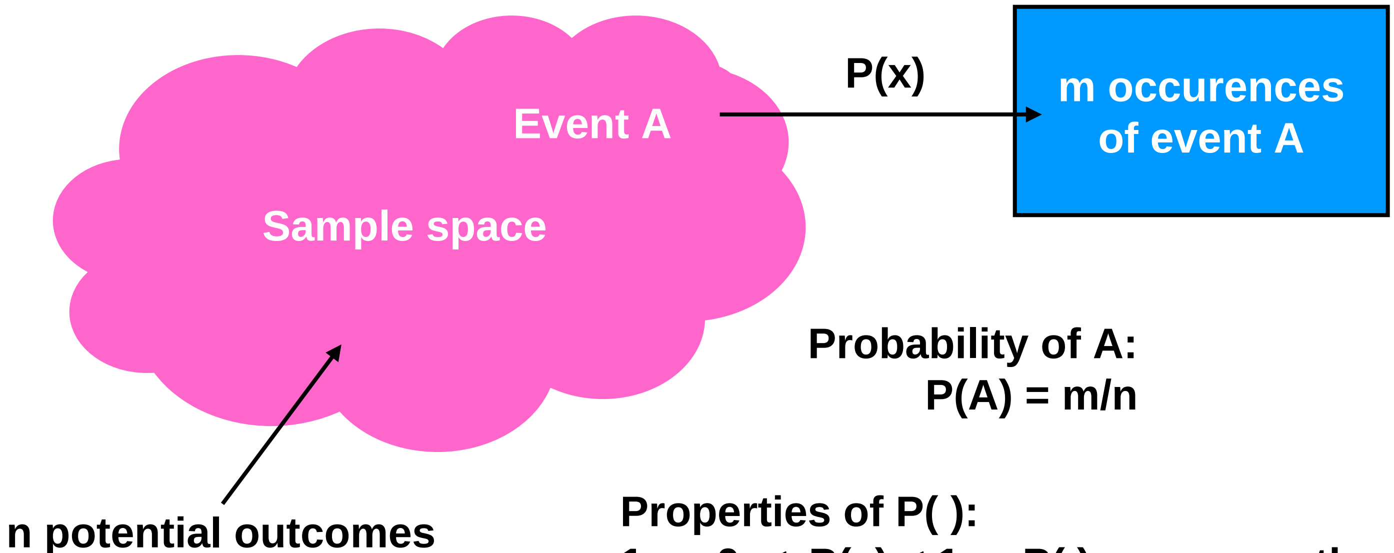
Probability of A:

P(A) = m/n

Properties of P( ):

1. **$0 \leq P(x) \leq 1$ - P( ) measures the proportion of A**

Probability Measures



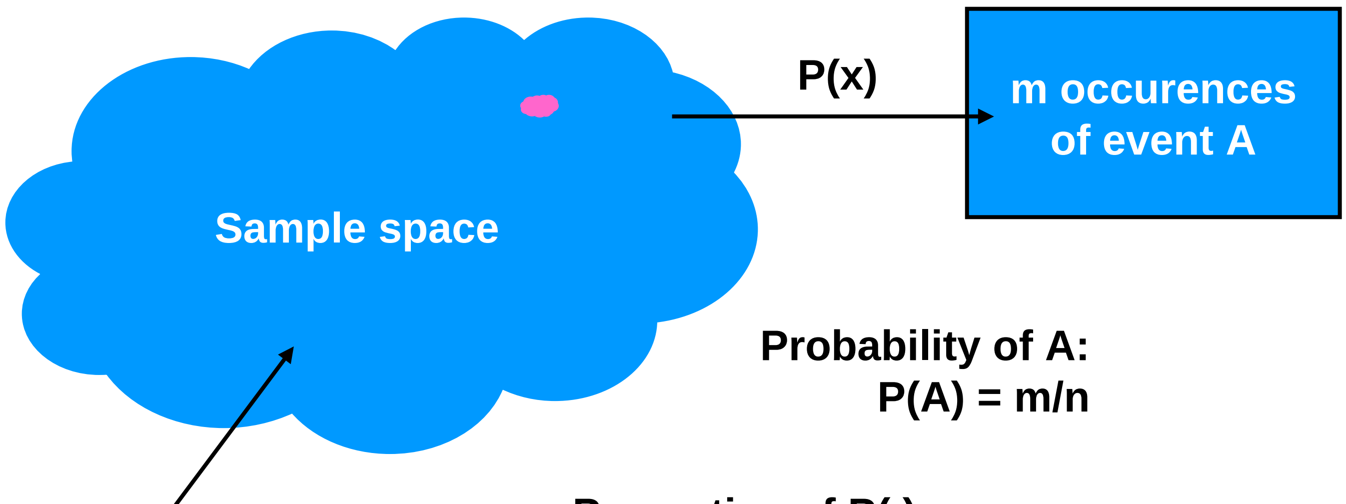
Probability of A:

P(A) = m/n

Properties of P( ):

1. $0 \leq P(x) \leq 1$ - P( ) measures the proportion of A
2. If A is certain, P(A) = 1

Probability Measures

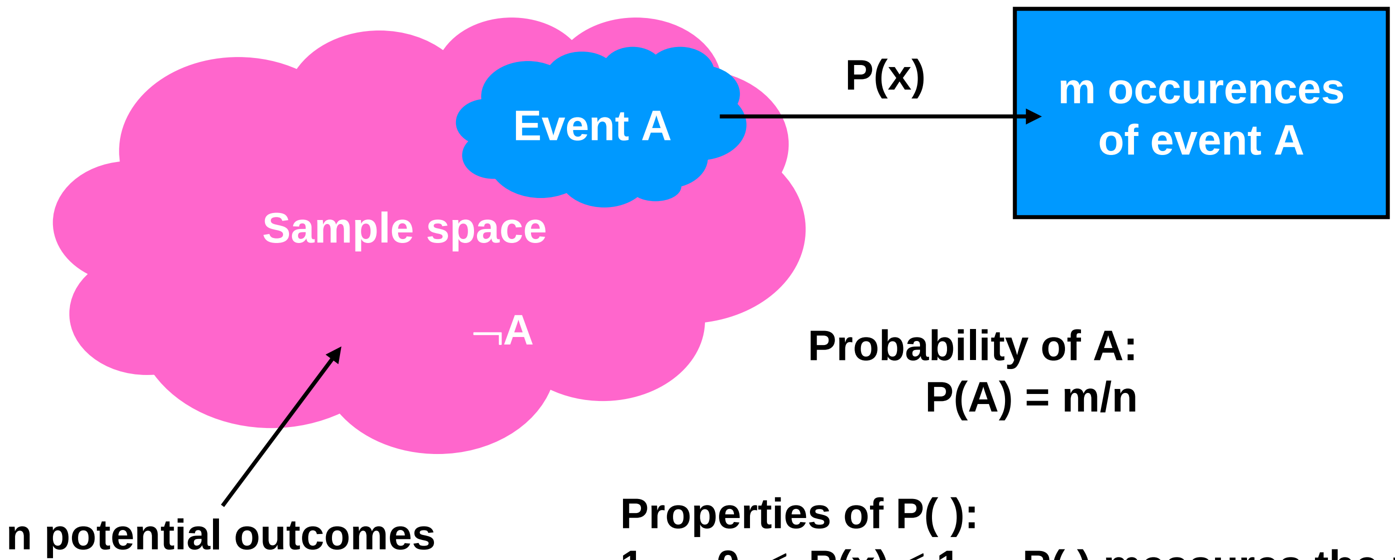


Probability of A:
P(A) = m/n

Properties of P( ):

1. $0 \leq P(x) \leq 1$ **- P( ) measures the proportion of A**
2. **If A is certain, P(A) = 1**
3. **If A is impossible, P(A) = 0**

Probability Measures

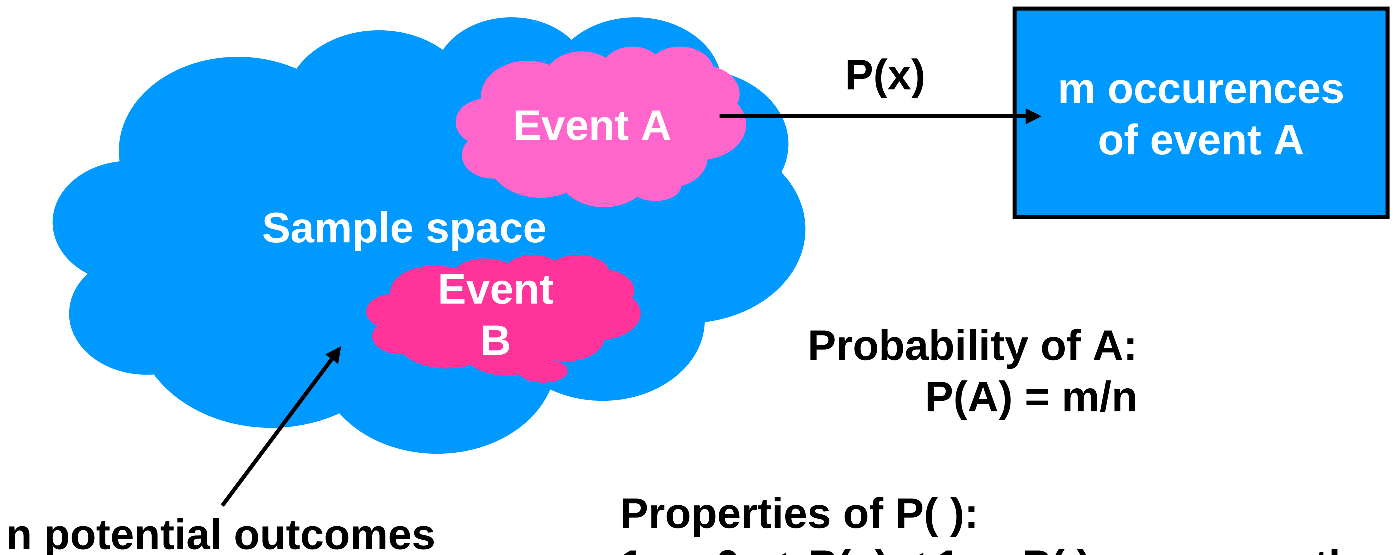


Probability of A:
P(A) = m/n

Properties of P( ):

1. **0 ≤ P(x) ≤ 1 - P( ) measures the proportion of A**
2. **If A is certain, P(A) = 1**
3. **If A is impossible, P(A) = 0**
4. **¬A is the complement of A. P(¬A) = 1 – P(A)**

Probability Measures

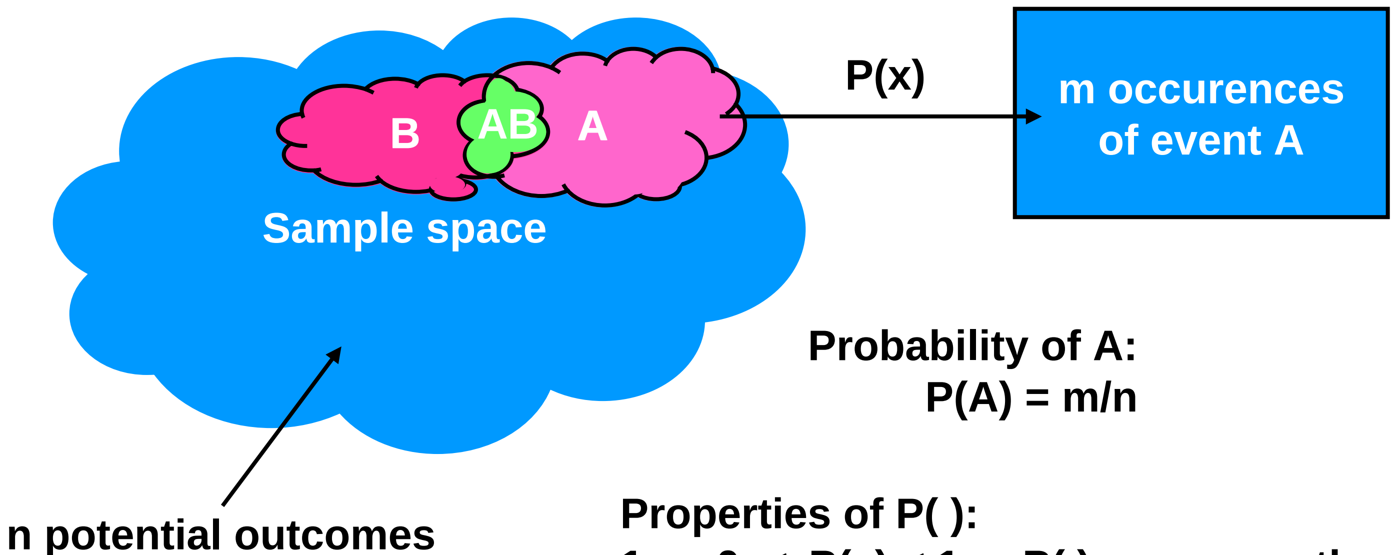


Probability of A:
P(A) = m/n

Properties of P( ):

1. **$0 \leq P(x) \leq 1$ - P( ) measures the proportion of A**
2. **If A is certain, P(A) = 1**
3. **If A is impossible, P(A) = 0**
4. **$\neg$A is the complement of A. $P(\neg A) = 1 - P(A)$**
5. **If A and B are mutually exclusive,**
   P(A or B) = P(A) + P(B)

Probability Measures

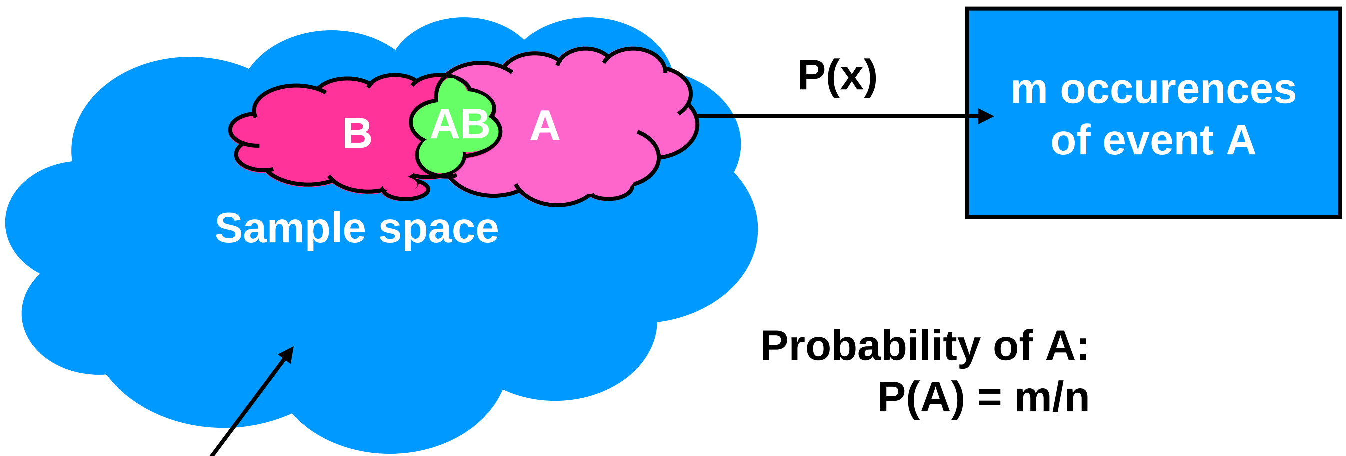


Probability of A:
P(A) = m/n

Properties of P( ):

1. $0 \leq P(x) \leq 1$ - P( ) measures the proportion of A
2. If A is certain, P(A) = 1
3. If A is impossible, P(A) = 0
4. $\neg A$ is the complement of A. $P(\neg A) = 1 - P(A)$
5. If A and B are mutually exclusive,
   P(A or B) = P(A) + P(B)
6. If A and B are independent events,
   P(AB) = P(A)P(B)

Probability Measures



n potential outcomes

Probability of A:
P(A) = m/n

Properties of P( ):

1. $0 \leq P(x) \leq 1$ - P( ) measures the proportion of A
2. If A is certain, P(A) = 1
3. If A is impossible, P(A) = 0
4. $\neg A$ is the complement of A. $P(\neg A) = 1 - P(A)$
5. If A and B are mutually exclusive,
   P(A or B) = P(A) + P(B)
6. If A and B are independent events,
   P(AB) = P(A)P(B)
7. Probability of either A or B (or both):
   P(Union of A, B) =
   $P(A \cup B) = P(A) + P(B) - P(AB)$

Next time

- More On Statistical Analysis of Experimental Data

Homework 7

- Problems 6.2, 6.4, 6.6