Bioeng 3070/5070

App Math/Stats for Bioengineer
Preliminaries

• Class Web Site: (Wiki Based)
  – http://www.bioen.utah.edu/wiki

• Homework assignments every week. With questions from the text. Towards the end of the semester Matlab based hands on data analysis.

• Homework Assignments (25%), Two Mid Term Exams (20%, 20%), Comprehensive Final Exam (30%), Class participation 5%

• Office: Web 3692.

• E-mail: sjoshi@sci.utah.edu

• Office hours: Mondays 2:00-3:30pm- open door policy (You can catch me anytime I am free or set up an appointment)
What is Probability and Statistics

• Formal definition: **Statistics** is a **mathematical science** pertaining to the collection, analysis, interpretation or explanation, and presentation of data.

• Probability theory: A mathematical axiomatic theory to describe and predict averages of mass phenomena occurring sequentially or simultaneously. Eg:
  – If an experiment is performed $n$ times and the event $A$ occurs $n_a$ times, then, with a high degree of certainty, the relative frequency $n_a/n$ is close to $P(A)$, provided that $n$ is sufficiently large.
Probability and Statistics

Probability: I know what’s in the bucket and I pick at random what’s in my hand.

Statistics: I know what’s in my hand, what can I say about what’s in the bucket.
• The Game Show Puzzle:
  – Imagine that you are a contestant on a television game show. You are shown three large doors. Behind one of the doors is a new car, and behind each of the other two is a goat. To win the car, you simply have to choose which door it is behind. When you choose a door, the host of the show opens one of the doors you have not chosen, and shows you that there is a goat behind it. You are then given a choice; you may stick with your original choice, or you may switch to the remaining closed door.
  – What should you do to maximize your chances of winning the car?
• Warning: Statistics and Probability can be very tricky.
  – Example 1: “EVERYONE knows men are promiscuous by nature. ….Surveys bear this out. In study after study and in country after country, on average men report more, often many more, sexual partners than women. “ New York Time, The Myth, the Math, the Sex.
  – CDC Survey:

<table>
<thead>
<tr>
<th>Partners</th>
<th>0-1</th>
<th>2-6</th>
<th>7-14</th>
<th>15 or more partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>16.6</td>
<td>33.8</td>
<td>20.7</td>
<td>28.9</td>
</tr>
<tr>
<td>Women</td>
<td>25.0</td>
<td>44.3</td>
<td>21.3</td>
<td>9.4</td>
</tr>
</tbody>
</table>
Simple Descriptive Statistics

• Arithmetic Mean: Sum of all observations divided by the number of observations:

\[ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \]

Properties of Arithmetic Mean: It is Linear in the observations. If

\[ y_i = ax_i + b, \ i = 1, \ldots, n \]

\[ \bar{y} = a \bar{x} + b \]
The Sample Median

• The “central” most point in the observations. Equal number of observations on both sides of the Sample Median.

• The sample median:
  1. \((\frac{n+1}{2})\)th Observation if \(n\) is odd
  2. The average of \((\frac{n}{2})\)th and \(\left(\frac{n}{2} + 1\right)\)th observation if \(n\) is even
Relationship between the Mean and the Median

- Symmetric distribution: Mean = Median
- Positively skewed distributions: Mean is greater than the Median.
- Negatively skewed distribution: Mean smaller than the median.

- The median is robust to Outliers. The Mean is very sensitive to outliers.
Mode

• The Mode is the most frequently occurring value among all the observations in a sample.
  – Mode might not be unique:
    • Distribution with just one mode is called **unimodal**
    • Distribution with two modes is called bimodal.
    • Distribution with three modes is called trimodal.
Geometric Mean

• The geometric mean is used for measurements that are meant to be multiplied together or are exponential in nature, for example concentrations and investment returns.

• It is defined as:

\[
\left( \prod_{i=1}^{n} x_i \right)^{1/n} = \exp \left( \frac{1}{n} \sum_{i=1}^{n} \log(x_i) \right)
\]
Geometric and Arithmetic Means

• Arithmetic Means: "if all the quantities had the same value, what would that value have to be in order to achieve the same total?"

• Geometric Mean: "if all the quantities had the same value, what would that value have to be in order to achieve the same product?"
Example of Geometric Mean

• For example, suppose you have an investment which earns 10% the first year, 50% the second year, and 30% the third year. What is its average rate of return? It is not the arithmetic mean, because what these numbers mean is that on the first year your investment was multiplied (not added to) by 1.10, on the second year it was multiplied by 1.60, and the third year it was multiplied by 1.20. The relevant quantity is the geometric mean of these three numbers.

• The question about finding the average rate of return can be rephrased as: "by what constant factor would your investment need to be multiplied by each year in order to achieve the same effect as multiplying by 1.10 one year, 1.60 the next, and 1.20 the third?" The answer is the geometric mean. If you calculate this geometric mean you get approximately 1.283, so the average rate of return is about 28% (not 30% which is what the arithmetic mean of 10%, 60%, and 20% would give you).