Week 3

D. Alex Hughes,
dhughes@ucsd.edu

October 6, 2010
A variable is an empirical measurement of a characteristic. (Remember the characteristics that were the essence of the concepts we were interested in?)

Variables have two features:

» Name – what we use to refer to the variable (Age, Married?, var_1)

» Values – what we are interested in. **Note:** is must have 2 values. Otherwise, it’s a constant!
Describing Variables

Central Tendency:

- **Mean:** \( \left( \frac{1}{N} \right) \sum_{i=1}^{N} x_i \)
- **Median:** Choose \( \left( \frac{N}{2} \right) \) from Ordered Set
- **Mode:** Maximally occurring observation. Useful for nominal-level variables.

Dispersion:

- **Ordinal:** Comparison of Median and Mode
- **Interval and Ratio:** Variance

\[
\sum_{i=1}^{N} (x_i - \bar{x})^2
\]

**Variance:** \( \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \)

**Standard Deviation:** \( \sigma = \left( \sigma^2 \right)^{\frac{1}{2}} \)
Wednesday Lecture

How many times did you guys see these graphs?

Distribution of Results

Value
Density of Results
20 30 40 50 60 70 80
0.000 0.005 0.010 0.015 0.020 0.025 0.030 0.035
Figure: Random Draws from a Binomial Distribution
<table>
<thead>
<tr>
<th>$x_i$</th>
<th>$x_i - \bar{x}$</th>
<th>$(x_i - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>-3</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Sum</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

$\sigma^2 = \frac{55}{8}$

$\sigma = \left(\frac{55}{8}\right)^{\frac{1}{2}}$

2.73