Section 3.2

# Linear Regression

PART V

Least Squares Regression Line: $\hat{y} = a + bx$

$b = \text{slope} = r \cdot \frac{s_y}{s_x}$, $a = y\text{-intercept} = \bar{y} - b\bar{x}$.

Developing a Computational formula for $b$ = slope,

$$b = r \cdot \frac{s_y}{s_x} = \frac{\Sigma xy - \frac{1}{n}(\Sigma x)(\Sigma y)}{\sqrt{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2}\sqrt{\Sigma y^2 - \frac{1}{n}(\Sigma y)^2}} \cdot \frac{\sqrt{\Sigma y^2 - \frac{1}{n}(\Sigma y)^2}\cdot\frac{1}{\sqrt{n-1}}}{\sqrt{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2}\cdot\frac{1}{\sqrt{n-1}}}$$

$$b = \frac{\Sigma xy - \frac{1}{n}(\Sigma x)(\Sigma y)}{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2} \leftarrow \text{sample slope of LSR line}$$

**Computational Formulas: LSR line: $\hat{y} = a + bx$**

**where $b = \frac{\Sigma xy - \frac{1}{n}(\Sigma x)(\Sigma y)}{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2}$, and $a = \bar{y} - b\bar{x}$.**

Automatic Weapons (Registered) vs. Murder Rate (cities)

| $x$ = auto weapons | $y$ = murder rate | $x^2$ | $y^2$ | $xy$ |
|---|---|---|---|---|
| 11.6 | 13.1 | 134.56 | 171.61 | 151.96 |
| 8.3 | 10.6 | 68.89 | 112.36 | 87.98 |
| 3.6 | 10.1 | 12.96 | 102.01 | 36.36 |
| 0.6 | 4.4 | 0.36 | 19.36 | 2.64 |
| 6.9 | 11.5 | 47.61 | 132.25 | 79.35 |
| 2.5 | 6.6 | 6.25 | 43.56 | 16.50 |
| 2.4 | 3.6 | 5.76 | 12.96 | 8.64 |
| 2.6 | 5.3 | 6.76 | 28.09 | 13.78 |
| 38.5 | 65.2 | 283.15 | 622.20 | 397.21 |
| $\uparrow \Sigma x$ | $\uparrow \Sigma y$ | $\uparrow \Sigma x^2$ | $\uparrow \Sigma y^2$ | $\uparrow \Sigma xy$. |

Sample size: $n = 8$

- Slope

$$b = \frac{\Sigma xy - \frac{1}{n}(\Sigma x)(\Sigma y)}{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2} = \frac{397.21 - \frac{1}{8}\cdot 38.5 \cdot 65.2}{283.15 - \frac{1}{8}\cdot 38.5^2} \approx 0.8525.$$

- $y$-intercept $a = \bar{y} - b\bar{x}$

$$a = \frac{65.2}{8} - 0.8525 \cdot \frac{38.5}{8} \approx 4.0473$$

LSR Line $\hat{y} = a + bx$; $\boxed{\hat{y} = 4.0473 + 0.8525 \cdot x}$