Chapter 3

"Class, I've got a lot of material to cover, so to save time I won't be using vowels today. Nw lts bgn, pls trn t pg 122."

Measures of Association Between Two Variables

- ✓ Covariance
- ✓ Pearson Product Moment Correlation Coefficient

Covariance – conceptual formula

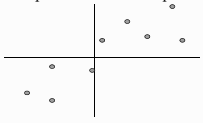
Sample Covariance

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

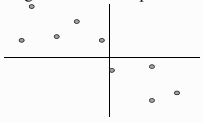
Sample variance

$$s_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Positive values indicate a positive relationship.



Negative values indicate a negative relationship.



- ✓ Covariance - the descriptive measure of the relationship between two variables
  - Population

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N} \qquad \sigma_{xy} = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{N^2}$$

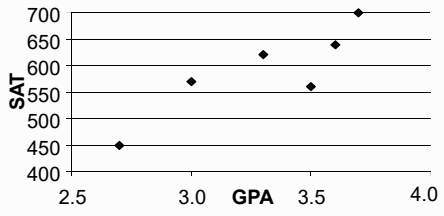
  - Sample

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} \qquad s_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n(n-1)}$$

Covariance

GPA	x_i^2	SAT	y_i^2	$x_i y_i$
2.7		450		
3.5		560		
3.7		700		
3.3		620		
3.6		640		
3.0		570		
19.8		3540		

Scatter Graph



Covariance

Page 117 - Problem 47

GPA	x_i^2	SAT	y_i^2	$x_i y_i$
2.7	7.3	450	202500	1215
3.5	12.3	560	313600	1960
3.7	13.7	700	490000	2590
3.3	10.9	620	384400	2046
3.6	13.0	640	409600	2304
3.0	9.0	570	324900	1710
19.8	66.1	3540	2125000	11825

$s_x = .3847_{GPA}$ $s_y = 85.323_{SAT}$

Covariance

$$s_{xy} = \frac{6(11825) - (19.8)(3540)}{6(5)} = 28.6_{GPA \cdot SAT}$$

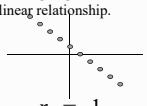
Correlation Coefficient

Pearson Product Moment Correlation Coefficient

Sample Correlation Coefficient $r_{xy} = \frac{s_{xy}}{s_x s_y}$

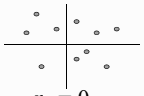
The coefficient can take on values between -1 and +1: $-1 \le r_{xy} \le 1$

Values near **-1** indicate a strong negative linear relationship.



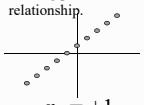
$r_{xy} = -1$

Values near **0** indicate no linear relationship.



$r_{xy} = 0$

Values near **+1** indicate a strong positive linear relationship.



$r_{xy} = +1$

Correlation Coefficient

- Pearson Product Moment Correlation Coefficient (sample correlation coefficient)
  - Population $\rho_{xy} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$ $\quad -1 \le \rho_{xy} \le 1$
  - Sample $r_{xy} = \dfrac{s_{xy}}{s_x s_y}$ $\quad -1 \le r_{xy} \le 1$

Correlation Coefficient

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

$$r_{xy} = \frac{28.6_{GPA \cdot SAT}}{(.3847_{GPA})(85.323_{SAT})} = \mathbf{.8713}$$