

1. In some cases, measures of association for categorical variables are useful even when the underlying variables are continuous if the question being addressed can be expressed in terms of variables that take only two values. For instance, suppose we are interested in the relationship between Tom Brady's success in passing in a given game and the pressure exerted by the defense's pass rush. Although there are many variables that could be used to address this issue in a formal analysis, here we consider the relationship between touch-down passes, denoted by T , and sacks, denoted by S .

Table 1: Sacks and touch-down passes for Brady in the 2009-2012 seasons

		Touch-down Passes		
		0-2	3 or more	Total
2*Sacks	0-2	29	22	51
	3 or more	18	2	20
Total		47	24	

- (a) (8 points) Compute the correlation coefficient. Interpret this number. Is this correlation significant?

The correlation is given by

$$r = \frac{29 \times 2 - 18 \times 22}{\sqrt{51 \times 20 \times 47 \times 24}} = -0.315$$

and the margin of error is

$$ME = 2\sqrt{\frac{1 - r^2}{n}} = 2\sqrt{\frac{1 - (-0.315)^2}{71}} = 0.225$$

The relationship between Sacks and touch-down passes is negative. Since the margin of error is less than the correlation in absolute value, the correlation is significant.

- (b) (5 points) Compute the α , the cross-product ratio. Interpret this number.

The cross-product ratio is given by

$$\alpha = \frac{ad}{bc} = \frac{29 \times 2}{18 \times 22} = 0.147$$

The odds of Brady throwing three or more touch-down passes is about 7 times greater if he is sacked less than three times than if he is sacked three or more times.

- (c) (5 points) Compute the Q , Yule's Q . Interpret this number.

The Yule's Q is given by

$$Q = \frac{\alpha - 1}{\alpha + 1} = \frac{0.147 - 1}{0.147 + 1} = -0.744$$

It follows that the relationship between Sacks and touch-downs is negative.