Section 9.3

Theorem: Let r be the sample correlation coefficient computed using data pairs (x, y). We use the null hypothesis

$$H_0\text{: } x \text{ and } y \text{ have no linear correlation, so } \rho = 0$$

The alternate hypothesis may be

$$H_1\text{: } \rho > 0 \quad \text{or} \quad H_1\text{: } \rho < 0 \quad \text{or} \quad H_1\text{: } \rho \neq 0$$

The conversion of r to a Student's t distribution is

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \text{ with df} = n - 2$$

where n is the number of sample data pairs (x, y) and $n \geq 3$.

Exercise 1. The weights (in pounds) of eight vehicles and the variability of their braking distances (in feet) when stopping on a wet surface are shown in the table. At $\alpha = 0.05$, can you conclude that there is a significant linear correlation between vehicle weight and variability in braking distance on a wet surface?

Weight, x	5800	5340	6500	4800	5940	5600	5100	5850
Variability in braking distances, y	2.92	2.40	4.09	1.72	2.88	2.53	2.32	2.78

Exercise 2. Let x be a random variable that represents the batting average of a professional baseball player. Let y be a random variable that represents the percentage of strikeouts of a professional baseball player. A random sample of $n = 6$ professional baseball players gave the following information (Reference: *The Baseball Encyclopedia*, Macmillan).

x	0.328	0.290	0.340	0.248	0.367	0.269
y	3.2	7.6	4.0	8.6	3.1	11.1

Use a 5% level of significance to test the claim that $\rho \neq 0$. (#8)

Class Exercise 1. Wolf packs tend to be large extended family groups that have a well-defined hunting territory. Wolves not in the pack are driven out of the territory or killed. In ecologically similar regions, is the size of an extended wolf pack related to size of hunting region? Using radio collars on wolves, the size of the hunting region can be estimated for a given pack of wolves. Let x represent the number of wolves in an extended pack and y represent the size of the hunting region in $km^2/1000$. From Denali National Park we have the following data

x wolves	26	37	22	69	98
y $km^2/1000$	7.38	12.13	8.18	15.36	16.81

Use a 1% level of significance to test the claim $\rho > 0$. **Answer: p-value = 0.0086**

Class Exercise 2. The number of hours 13 students spent studying for a test and their scores on that test are shown in the table. Is there enough evidence to conclude that there is a significant linear correlation between the data? Use $\alpha = 0.01$.

Hours spent studying, x	0	1	2	4	4	5	5	5	6	6	7	7	8
Test score, y	40	41	51	48	64	69	73	75	68	93	84	90	95

Answer: p-value $= 7.08 \cdot 10^{-6}$

Class Exercise 3. The following table shows the number of crimes reported (in millions) and the number of arrests reported (in millions) by the U.S. Department of Justice for 14 years. At $\alpha = 0.05$, can you conclude that there is a significant linear correlation between the number of crimes and the number of states?

Crimes, x	1.66	1.65	1.60	1.55	1.44	1.40	1.32	1.23	1.22	1.23	1.22	1.18	1.16	1.19
Arrest, y	0.72	0.72	0.78	0.80	0.73	0.72	0.68	0.64	0.63	0.63	0.62	0.60	0.59	0.60

Answer: p-value $= 2.31 \cdot 10^{-5}$

Homework

C Problems

Section 9.3: 7(a), 7(b), 9(a), 9(b), 11(a), 11(b)

B Problems

Section 9.3: 1

A Problems

None