Hypothesis Test on Population Correlation Coefficient, $\rho$

The sample corr. coeff., $r$, estimates the pop. corr. coeff., $\rho$ (Rho)

We can perform a test of hypothesis on $\rho$

Ex: (cont. from sec. 11.1)

$H_0: \rho = 0$  (no correlation between $X$ and $Y$)

$H_a: \rho \neq 0$  (is a correlation between $X$ and $Y$)
\[ \alpha = .05 \]

**Test Statistic:**

\[ T = r \times \sqrt{\frac{n-2}{1-r^2}} \]

\[ = 0.926 \times \sqrt{\frac{8-2}{1-0.926^2}} \]

\[ = 0.926 \times \sqrt{\frac{6}{0.1425}} \]

\[ = 0.926 \times \sqrt{42.0982} \]

\[ = 6.01 \]
P-value of test statistic:

Since $H_a: \rho \neq 0$ (2-sided $H_a$)

$p$-value = $2 \times P(T > |\text{test stat. value}|)$

= $2 \times P(T > 6.01)$

Use table F with $df = n-2 = 6$

we see $P(T > 3.71) = 0.005$

So $P(T > 6.01) < 0.005$ and $p$-value < $2 \times 0.005 = 0.010$

Decision: Reject $H_0$ if $p$-value $\leq \alpha$

Since $p$-value $< 0.010 \leq 0.05$

we reject $H_0$. 

Conclude: There does seem to be a correlation between "number of hours of TV per day" and cholesterol level, at the 5% significance level.

Assumptions:
Each variable is normally distributed.
The Connection Between Correlation and Regression

Consider a test on $\beta$:

$H_0: \beta = 0$

$H_a: \beta \neq 0$

If we fail to reject $H_0$, this indicates that our regression line is (approximately) horizontal.

This means that the X's provide little, if any,
value in predicting the y's, i.e., the predicted y would be (nearly) the same for all values of x.

This corresponds to saying that variables X and Y are uncorrelated.

In fact, were we to conduct the test

\[ H_0 : \rho = 0 \]
\[ H_a : \rho \neq 0 \]
we would fail to reject \( H_0 \).

Similarly, if we test

\[ H_0: \beta = 0 \]
\[ H_a: \beta \neq 0 \]

and reject \( H_0 \), we would also reject \( H_0 \) in the test

\[ H_0: \rho = 0 \]
\[ H_a: \rho \neq 0 \]
Look at the "TV hours" and "cholesterol level" examples. In Sec 11.1 we tested

\[ H_0: \rho = 0 \]
\[ H_a: \rho \neq 0 \]

at \( \alpha = 0.05 \)

\[ T = 6.01 \quad \text{(test statistic)} \]

\[ p\text{-value} = 2 \times P(T \geq 6.01) < 0.010 \]

and we rejected \( H_0 \)
In Sec 11.1 we tested

\[ H_0: \beta = 0 \]
\[ H_a: \beta \neq 0 \]

at \( \alpha = 0.05 \)

\[ T = 6.03 \quad \text{(test-stat)} \]

\[ p-value = 2 \times P(T \geq |test\ stat\ value|) \]
\[ = 2 \times P(T \geq 6.01) < 0.010 \]

and we reject \( H_0 \).
Note that in both tests,

1) the same p-values
2) the test-statistic values are identical (except for roundoff error)
3) same decision (Reject $H_0$)