

An agronomist expects that, on average, high bush blueberry production will be negatively associated with cloud cover. The agronomist obtains records of cloud cover and berry production. The observed correlation is $r = -0.40$ based on 15 years. Test whether correlation is significantly less than zero (one-tailed test).

For small sample sizes the statistic t_s is normally distributed.

$$t_s = (z-0) (n-3)^{1/2} \quad \text{where} \quad z = (0.5) \ln\left(\frac{1+r}{1-r}\right)$$

Thus we can use the normal distribution to calculate p-values for t_s . Here is the cumulative distribution function for negative values of t_s , at values of r ranging from 0 to -0.9

```
MTB > set into c1
DATA> 0 -.1 -.2 -.3 -.4 -.5 -.6 -.7 -.8 -.9
DATA> end
MTB > let c2 = 0.5*log((1+c1)/(1-c1))*sqrt(15-3)

MTB > cdf c2;
SUBC> normal 0 1.
    0.0000    0.5000
   -0.3476    0.3641
   -0.7023    0.2413
   -1.0722    0.1418
   -1.4676    0.0711
   -1.9029    0.0285
   -2.4011    0.0082
   -3.0044    0.0013
   -3.8057    0.0001
   -5.0999    0.0000
```

column 1 of the output is the normal score (z) for t_s values of r ranging from 0 to -0.9
column 2 of the output is the p-value corresponding to several negative values of z and hence the t_s statistic.

What is the probability of obtaining a normal score of -1.9 or less? 0.0285

The normal distribution is symmetrical.

What is the probability of obtaining a normal score of 1.9 or more? 0.0285

What is the value of t_s when $r = 0$? $t_s = 0$ when $r = 0$

Be sure to state null and alternative hypotheses concerning r,

$$\begin{aligned} H_A: r < 0 & \quad \text{or equivalently} \quad H_A: t_s < 0 \\ H_0: r \geq 0 & \quad \text{or equivalently} \quad H_0: t_s \geq 0 \end{aligned}$$

state your significance criterion, $\alpha = 5\%$ (or $\alpha = 1\%$ or $\alpha = 10\%$)

calculate the t-statistic for the observed correlation ($r = -0.40$), -1.4676

and declare a decision.

Correlation not significant at 5%
 $r = -0.40, p = 0.071, n = 15$