

An agronomist expects that, on average, low 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 30 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(30-3)

MTB > cdf c2;
SUBC> normal 0 1.
    0.0000    0.5000
   -0.5214    0.3011
   -1.0534    0.1461
   -1.6083    0.0539
   -2.2013    0.0139
   -2.8543    0.0022
   -3.6017    0.0002
   -4.5066    0.0000
   -5.7086    0.0000
   -7.6499    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  $-3.6$  or less? \_\_\_\_\_

The normal distribution is symmetrical.

What is the probability of obtaining a normal score of  $3.6$  or more? \_\_\_\_\_

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

Be sure to state null and alternative hypotheses concerning  $r$ ,

state your significance criterion, \_\_\_\_\_

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

and declare a decision.