

Video 25 Tests of Significance

1. In the 1970s, statistician Ron Thisted did a statistical analysis of Shakespeare's vocabulary. Based on his analysis he created a computer program. What could his program tell you about a Shakespearean poem?

The program can determine the number of new words in a poem that Shakespeare did not use in any of his other writings.

2. In analyzing a poem to see whether or not it was authored by Shakespeare, Thisted set up a null hypothesis and an alternative hypothesis. State those hypotheses in words.

The null hypothesis was that Shakespeare was the author of the poem. The alternative hypothesis was that someone else wrote the poem.

3. What was the approximate distribution of the number of unique words per poem in Shakespeare's poems?

it is normally distributed with the mean $\mu = 7$ and standard deviation $\sigma = 2.6$.

4. Thisted observed 10 unique words in the newly discovered poem. Was that sufficient evidence to conclude that Shakespeare did not write the poem?

No. The value will be at least 10 unique words about 25% of the time when the poems were Shakespeare's.

5. Which is better evidence against the null hypothesis, a large p-value or a small p-value?

A small p-value