

Is the following data random?

16, 25, 52, 11, 38, 47, 12, 98, 4

**Example 17.5.3**

**Detecting Randomness of a Set of Numbers**

**Solution**

How do you test randomness with a numerical data set? Create a new data set comparing each value to the median value. To do this, substitute each value in the original data set with an A if it is above the median value, a B if it is below the median value, and eliminate any values that equal the median.

$H_0$ : The data is random.

$H_a$ : The data is not random.

Median = 25

16, 25, 52, 11, 38, 47, 12, 98, 4

B,  $\emptyset$ , A, B, A, A, B, A, B

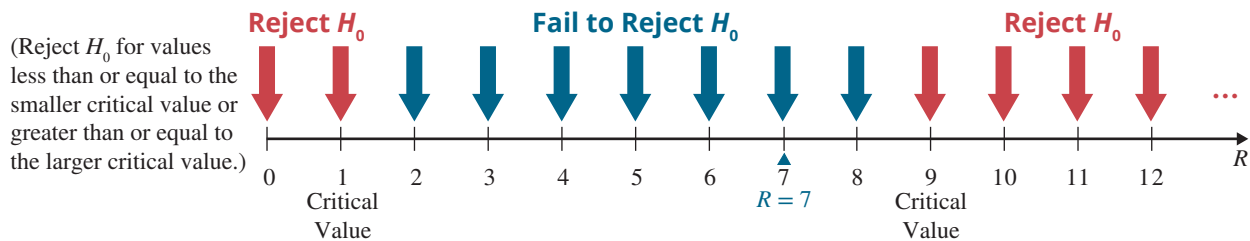
$m$  (the number of A's) = 4

$n$  (the number of B's) = 4

$R$  (the number of runs) = 7

critical values = 1, 9

Fail to reject  $H_0$ ; there is no evidence of nonrandomness.



## 17.5 Exercises

### Basic Concepts

1. Describe in your own words what is being tested with the runs test.
2. Consider the following sequence of 10 coin tosses.

H, H, T, T, H, H, H, T, T, H

Without performing any kind of test, do you believe this sequence is random? Explain why or why not.

3. What are the null and alternative hypotheses associated with the runs test?
4. What parameters need to be calculated in order to perform a runs test?
5. What is the rejection rule for a small sample runs test? How small is a *small* sample?
6. What is the rejection rule for a large sample runs test? How large is a *large* sample?
7. If a numerical set of data is under consideration, which parameter are the data points compared to in order to perform the runs test?

## Exercises

8. In the state of Tennessee, the number of deaths due to traffic accidents from 2010 to 2020 are shown in the following table. Use the runs test to examine non-randomness at the 0.05 level.

Traffic Fatalities in Tennessee, 2010-2020	
Year	Number of Traffic Fatalities
2010	1032
2011	937
2012	1014
2013	995
2014	963
2015	962
2016	1037
2017	1024
2018	1040
2019	1148
2020	1221

9. A sociologist designs a study that involves a procedure of selecting individuals randomly from an email list and then contacting them to determine if they own or rent their residence. The results are recorded in the order of phone calls (O = Own, R = Rent).

O O R R O R O R R O R R R R O R R R O O R R R O R

Does the sociologist have a random sequence of residential data at the 0.05 level?

10. A car tire manufacturer keeps track of the tires produced by one of the production lines. They observe the following sequence (D for defective items and N for non-defective items).

D D D N N D N D N D D D

Test the quality control manager's claim that there is no pattern in producing defective tires at the 0.05 level.

11. A marathon runner tries to run every day except when it is raining during the month of July. He observes the rainy (R) days and sunny (S) days to be able to predict the weather as follows.

S S S R R S S S R R R R S R S R R S S R S R S R R S R S S

Are the rainy days randomly scattered in the month of July at the 0.05 level?