

Is the iPod Random Shuffle Really Random?

In *The Guardian*, Steven Levy wrote about an interview with Steve Jobs, CEO of Apple, in which he presented Jobs with this dilemma: "I have a situation with my iPod. The shuffle function just doesn't seem random. Some artists come up way too much and some don't come up at all." According to Jeff Robbin, a head of the iTunes development team, "It is absolutely unequivocally random." Mathematician John Allen Paulos said that "[w]e often interpret and impose patterns on events that are random." Levy goes on to state that when we think that the iPod shuffle is not random, the problem is in our perceptions. Our minds perceive patterns and trends that don't really exist. We often hear runs of consecutive songs by the same artist and think that this is not random, but with true randomness, such consecutive runs are much more likely than we would expect.

The incorrect perception of nonrandomness caused Apple to introduce a "smart shuffle" feature in a new version of iTunes. This feature allows users to control multiple consecutive songs by the same artist. With this feature,

consecutive runs by the same artist would be avoided. According to Steve Jobs, "We're making it less random to make it feel more random."



From the display above we can see that there are 5 Eastern Conference winners and 10 Western Conference winners, and the number of runs is 8. We therefore have

$$n_1 = \text{number of Eastern Conference winners} = 5$$

$$n_2 = \text{number of Western Conference winners} = 10$$

$$G = \text{number of runs} = 8$$

Because $n_1 \leq 20$ and $n_2 \leq 20$ and the significance level is $\alpha = 0.05$, the test statistic is $G = 8$ (the number of runs), and we refer to Table A-10 to find the critical values of 3 and 12. Because $G = 8$ is neither less than or equal to the critical value of 3, nor is it greater than or equal to the critical value of 12, *we do not reject randomness*. There is not sufficient evidence to reject randomness in the sequence of winners. Randomness cannot be rejected.

Numerical Data: Randomness Above and Below the Mean or Median In Example 1 we tested for randomness in the sequence of data that clearly fit into two categories. We can also test for randomness in the way numerical data fluctuate above or below a mean or median. To test for randomness above and below the median, for example, use the sample data to find the value of the median, then replace each individual value with the letter A if it is *above* the median and replace it with B if it is *below* the median. Delete any values that are equal to the median. It is helpful to write the A's and B's directly above or below the numbers they represent because this makes checking easier and also reduces the chance of having the wrong number of letters. After finding the sequence of A and B letters, we can proceed to apply the runs test as described earlier. Economists use the runs test for randomness above and below the median to identify trends or cycles. An upward economic trend would contain a predominance of B's at the beginning and A's at the end, so the number of runs would be very small. A downward trend would have A's dominating at the beginning and B's at the end, with a small number of runs. A cyclical pattern would yield a sequence that systematically changes, so the number of runs would tend to be large.

Example 2 Testing Randomness Above and Below the Median

Listed below are the consecutive annual high values of the Dow Jones Industrial Average for the 1990s. Use a 0.05 significance level to test for randomness above and below the median.

| | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|--------|
| 3000 | 3169 | 3413 | 3794 | 3978 | 5216 | 6561 | 8259 | 9374 | 11,568 |
| ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ |
| B | B | B | B | B | A | A | A | A | A |

Solution

Requirement check (1) The data are arranged in order. (2) Each data value is categorized into one of two separate categories (below the median or above the median). The requirements are satisfied. ✓

The median of the listed sample values is 4597.0. We denote a value *below* the median of 4597.0 by B (below) and we denote a value *above* the median by A (above). If there had been any values equal to the median, they would have been deleted. The sequence of B's and A's is shown below the sample values. That