significant, and the null hypothesis is rejected. What is rejected is the notion that smoking does not significantly influence lung volume.

In configuring the smoking study, the statistician can ask, "Does smoking significantly reduce lung volume?" This involves using a one-sided P value. Alternatively, the statistician can ask, "Does smoking significantly change lung volume?" This involves using a two-sided P value. The value of the two-sided P value is twice 0.0082, that is, $P = 0.0164$.

Please note the negative sign in the above Z value, that is, $Z = -2.4$. For the purposes of plugging in the Z value and obtaining a probability, the negative sign can be ignored. In the words of Norman and Streiner (81), "what we do is ignore the sign, but keep it in our minds."

Pocock (82) provides additional examples from actual clinical trials. What is provided are numbers for plugging into the formula, $Z = (\overline{x}_1 - \overline{x}_0)/\left[\sqrt{(s_1^2/n_1) + (s_0^2/n_0)}\right]$, the calculated number for the Z value, instructions on plugging the Z value into the standard table, the value for P that was determined from this table, and instructions for interpreting the P value.

X. SUMMARY

To calculate the P value, the investigator starts with data, regarding an event of interest, from a first sample, and data, regarding the same event but from a second sample. From these data, the investigator then calculates the means, standard deviations, and the Z value. The Z value is then plugged into a standard table with numbers corresponding to areas under a normal distribution curve. In plugging in the Z value, the investigator then arrives at the P value. This particular routine is used where the distribution of values in the first sample (study drug) follows a normal distribution, and where the distribution in the second sample (control treatment) also follows a normal distribution. Where the distribution of values is not normal, that is, where the distribution is skewed or contains two peaks, the investigator should use a different statistical tool, that is, a statistical tool that is a nonparametric test.

Daniel (83) provides a flow chart (decision tree) for determining which statistical formula to use. The decision tree asks whether the population is normally distributed, if the sample is large or if the sample is small, and if the population variance (or standard deviation) is known or unknown. Depending on the answers, the researcher may need to use, or may prefer to use, the Z statistic, the t statistic, or a nonparametric test such as the Wilcoxon rank sum test.

XI. THEORY BEHIND THE Z VALUE AND THE TABLE OF AREAS IN TAIL OF THE STANDARD NORMAL DISTRIBUTION

In brief, calculating the Z value converts the raw data into a normalized value. The normalized value, when plugged into the standard table, provides an area under a curve that depicts the normal distribution. This area is, in effect, identical with the probability (P value). The goal of this chapter is to serve as a starting point in biostatistics, and to provide a reference point for use in navigating through textbooks on biostatistics.

[81]Norman GR, Streiner DL. Biostatistics. 3rd ed. Hamilton, Ontario: B.C. Decker, Inc.; 2008. p. 35.

[82]Pocock SJ. The simplest statistical test:how to check for a difference between treatments. Brit. Med. J. 2006;332:1256–58.

[83]Daniel WW. Biostatistics. 9th ed. Hoboken, NJ: John Wiley & Sons, Inc.; 2009. p. 176.