COMPARING THE DURABILITY OF TIRES

14. Brief Version of the Case Study

14.1 Problem Formulation
14.2 Study Design
14.3 Displaying and Describing the Data
14.4 Using the t-Tools
14.5 Using the Nonparametric Methods
14.6 Summary

14.1 Problem Formulation

A tire manufacturer is experimenting with a new technology, which hopefully will produce more durable tires. He has recently developed a new type of tire (brand A) that is supposed to be more durable than its currently marketed tire (brand B). In order to compare the durability of the two brands, an experiment involving twenty automobiles was carried out. On each automobile, an A tire was randomly assigned to one of the rear wheels and a B tire was assigned to the other rear wheel. Then the automobiles were operated for 20,000 km under normal circumstances.

One possible measure of durability is tread depth. The new technology is supposed to reduce the tire wear, or equivalently increase the tread depth of tires. In order to evaluate the effectiveness of the technology, the tread depth (expressed in 1/32 of an inch) was determined and recorded for each of the forty tires in our experiment. The data are available in the SPSS file tires.sav located on the FTP server. The instructions how to access the data on the server and download it to your floppy disk are available in *Downloading Data Files Using FTP* module.

The following is a description of the variables in the data file:

Column	Name of Variable	Description of Variable
1	Brand A	Tread Depth (1/32 in.)
2	Brand B	Tread Depth (1/32 in.)

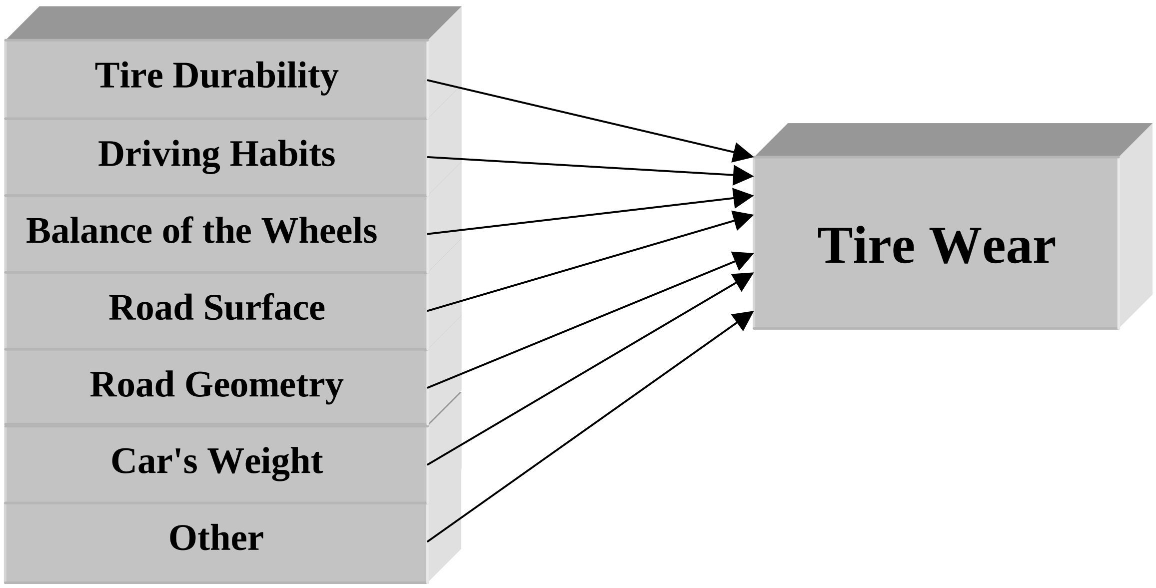
We will use SPSS to answer the following two questions using the data:

1. Do the data allow us to conclude that the brand A tires are more durable, on the average, than the brand B tires?

2. Estimate the change in the durability of tires after the new technology was implemented by means of a 95% confidence interval for the mean difference in the tread depth between brand A and brand B tires.

14.2 Study Design

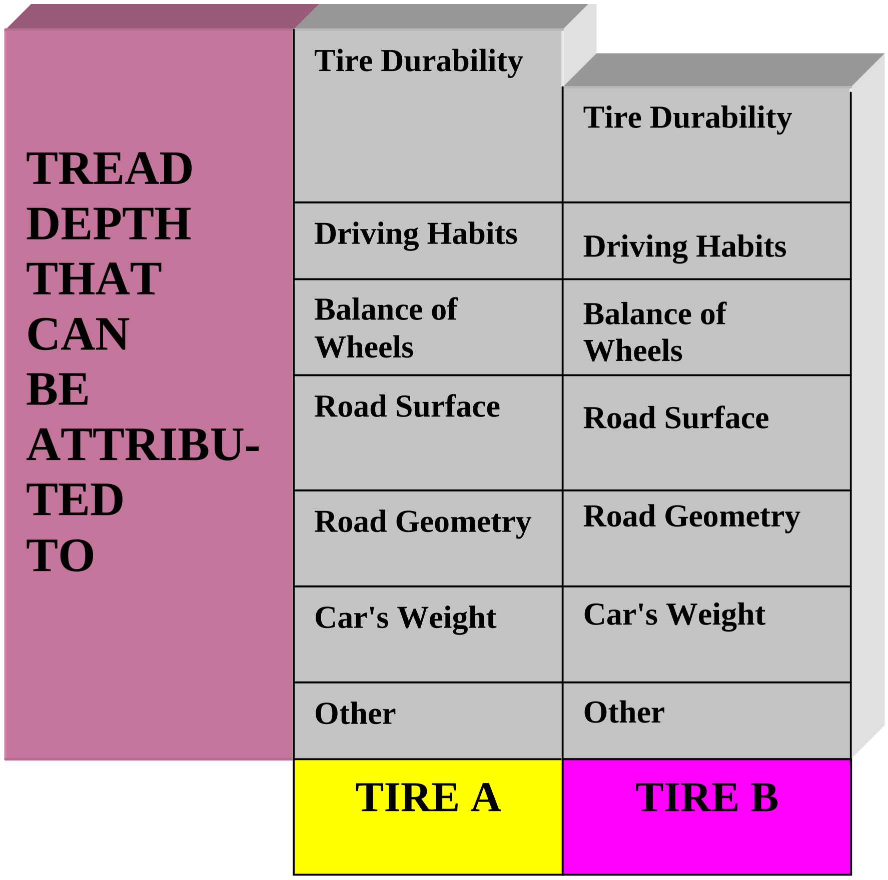
We start with identifying important variables affecting tire wear. The durability of a tire is determined by several factors, among them the quality of the material used to make the tire and the technology used in the manufacturing process. However, durability is only one of a multitude of factors affecting the tire wear. The tire wear is also greatly affected by the driving habits of the driver, the condition and types of roads driven on (road surface and road geometry), the balance of the wheels, the size and weight of the car, the age and condition of the car, climate, etc.



We want to compare the tire wear that can only be attributed to the tire durability but not to other factors.

In order to neutralize the other factors affecting tire wear, we will use a very special design called matched pairs design. The design in our experiment can be implemented in the following way. First 20 tires of brand A and 20 tires of brand B will be selected at random. Then twenty cars will be selected randomly and a pair of tires (one brand A, one brand B) will be installed on the rear wheel of each car. On each automobile, an A tire is randomly assigned to one of the rear wheels and a B tire is assigned to the other rear wheel. This procedure tends to eliminate the effects of the car-to-car variability and yields more information on the differences in the wearing quality of the two brands. Each driver will be driving the car 20,000 km under normal circumstances, and then the tread depth will be determined for each of them.

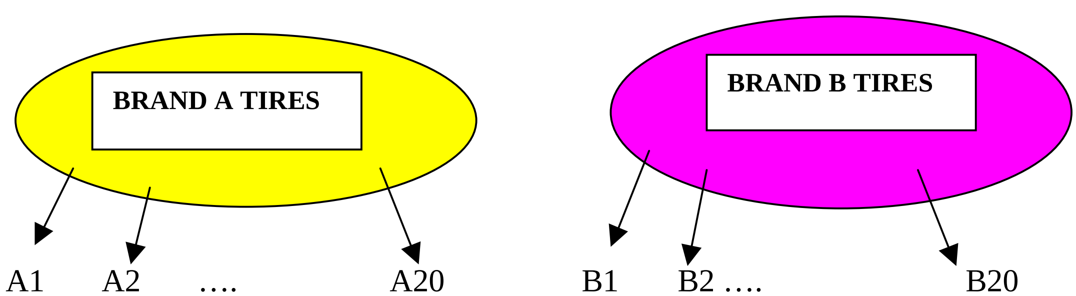
Observe that by assigning one A and one B tire to each of the twenty cars, the tires A and B were affected by the same driving habits (same driver), by the same balance of wheels (same car), the same conditions, and so on. Assigning two different brands of tires A and B to each particular automobile eliminates the effect of the car-to-car variability. Thus, any differences in the tire wear should be attributed to the differences in the durability of the two tires. This idea is illustrated on the following picture.



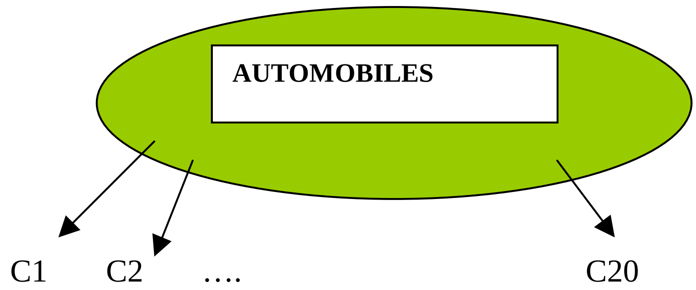
Since the various cars, drivers, and conditions are the same for each pair of tires mounted on the rear wheels of a particular automobile, it would make sense to calculate the difference D=A-B, where A and B are the tread depths of the tires A and B, respectively. Then D expresses the real difference in the durability of the two tires. Looking at the differences in the two measurements for each car neutralizes the variability among the cars.

Now we will describe how the randomization process can be carried out:

1. Random selection of 20 brand A tires and 20 brand B tires.



2. Random selection of 20 automobiles.



3. Random allocation of one A and one B tire to each of the 20 cars.

SELECTED BRAND A TIRES	A1	A2	……	…..	A20
CAR NUMBER	12	2			8

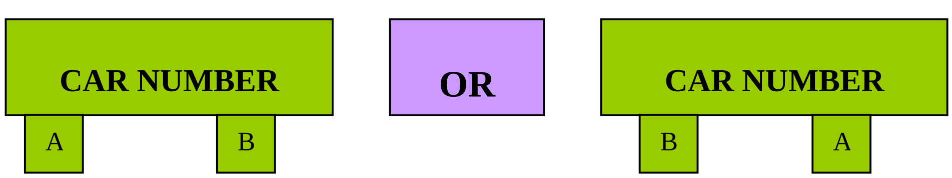
Read successive two-digit numbers between 1 and 20 from a table of random digits or use the random number generation feature in a computer software

SELECTED BRAND A TIRES	B1	B2	……	…..	B20
CAR NUMBER	4	19			6

The same random generation mechanism as described above

4. Random decision about the location of each tire on each car

   In the previous step, one brand A and one brand B tire is selected for each car. We have still to decide which tire goes to which wheel. Two possible solutions are:



In order to make the random decision, you can flip a coin. If the head occurs, assign the A tire to the left wheel, otherwise assign the B tire to the left wheel.

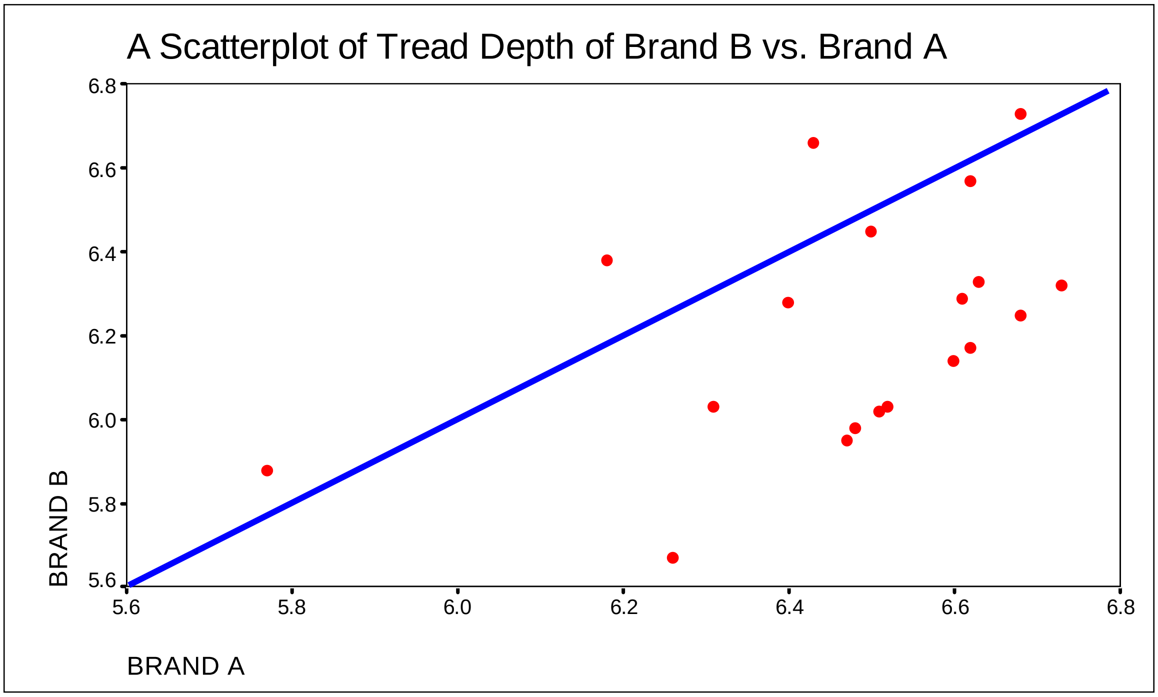
Our randomization process is complete. Now it is time to carry out the experiment and collect the data.

AUTOMOBILE	BRAND A	BRAND B	DIFFERENCE
1	8.46	8.05	0.41
2	6.47	5.95	0.52
3	6.63	6.33	0.30
4	5.98	5.52	0.46
5	6.61	6.29	0.32
6	6.06	5.40	0.66
7	5.77	5.88	-0.11
8	6.62	6.17	0.45
9	5.32	5.55	-0.23
10	6.31	6.03	0.28
11	6.51	6.02	0.49
12	6.18	6.38	-0.20
13	5.95	5.83	0.12
14	6.62	6.57	0.05
15	6.68	6.25	0.43
16	6.52	6.03	0.49
17	6.68	6.73	-0.05
18	6.48	5.98	0.50
19	5.89	5.30	0.59
20	7.00	6.95	0.05
AVERAGE	6.44	6.16	0.28

14.3 Displaying and Describing the Data

We will compare the durability of the two brands of tires using both graphical and numerical tools.

The figure below gives such a scatterplot for the data in our experiment.



The 45-degree line in the plot represents the pairs with X equal to Y (the tread depths of brand A are equal to the tread depths of brand B). If the new tires are more durable, we expect that the tread depths of brand A tires be larger than the corresponding tread depths of brand B tires. Thus, most of the points should be below the X=Y line. As sixteen of the 20 pairs are in the region below the X=Y line, the plot supports the claim that the tread depth of brand A tends to be larger than the corresponding tread depth of brand B.

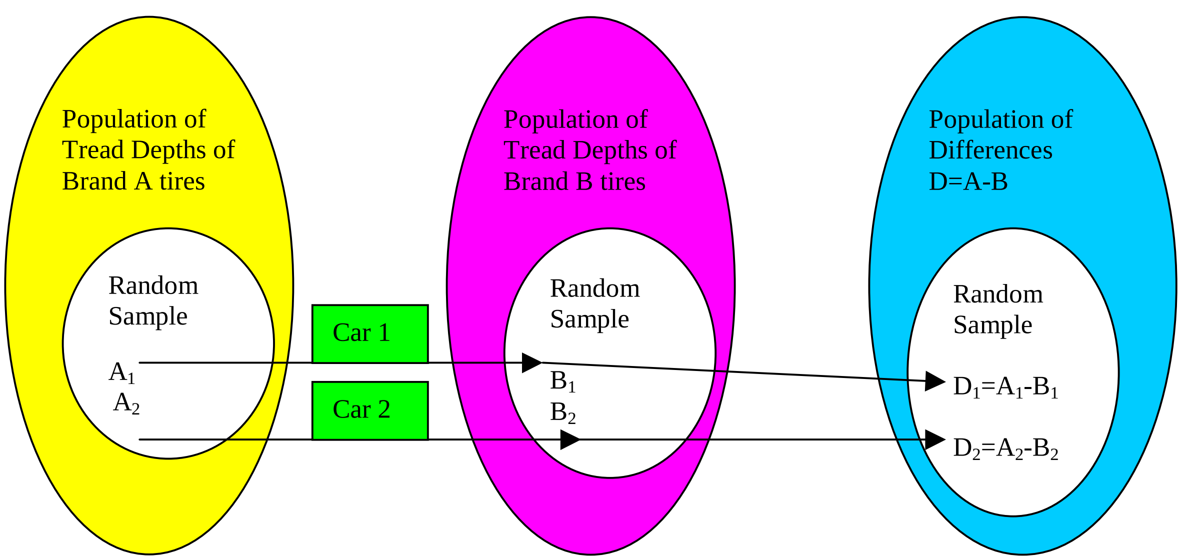
SPSS produces the following summary statistics for the distribution of differences.

	STATISTICS	DIFFERENCES D=A-B
MEASURES OF CENTER	MEAN	0.2765
	MEDIAN	0.3650
	5% TRIMMED MEAN	0.2833
	95% CI FOR MEAN	(0.1486, 0.4044)
MEASURES OF SPREAD	STANDARD DEV.	0.2734
	STANDARD ERROR	0.0611
	VARIANCE	0.0747
	IQR	0.4400
	MINIMUM	-0.2300
	MAXIMUM	0.6600
	RANGE	0.8900
MEASURES OF SHAPE	SKEWNESS	-0.5952
	ST. ERROR SKEWNESS	0.5121
	KURTOSIS	-0.9362
	ST. ERROR KURTOSIS	0.9924
COUNT		20

All displayed measures of center indicate that the typical difference in the tread depth between the brand A and brand B tires is positive indicating that the tread depth of brand A tires tends to exceed the tread depth of brand B tires.

14.4 Using the t-Tools

In order to isolate and measure the effects of tire durability on tire wear, we have used matched pairs design and restated the data as single-sample data by taking the differences within each pair. Hence, in order to make the inferences about the data, we will be using the statistical tools for single populations. In our case study, the population is the population of all differences within matched pairs.



Denote by μ_A and μ_B the population mean tread depth of all brand A and brand B tires, respectively. Then the difference between two population means μ_A and μ_B is equivalent to the mean of the paired differences. In other words, $\mu_D = \mu_A - \mu_B$, where D=A-B.

We can use the sample of 20 differences to make inferences about the mean of the population of differences, μ_D -which is equal to the difference μ_A-μ_B. Thus, our test becomes

$$H_0: \mu_D=0 \ (\mu_A-\mu_B=0) \text{ versus } H_a: \mu_D>0 \ (\mu_A-\mu_B>0).$$

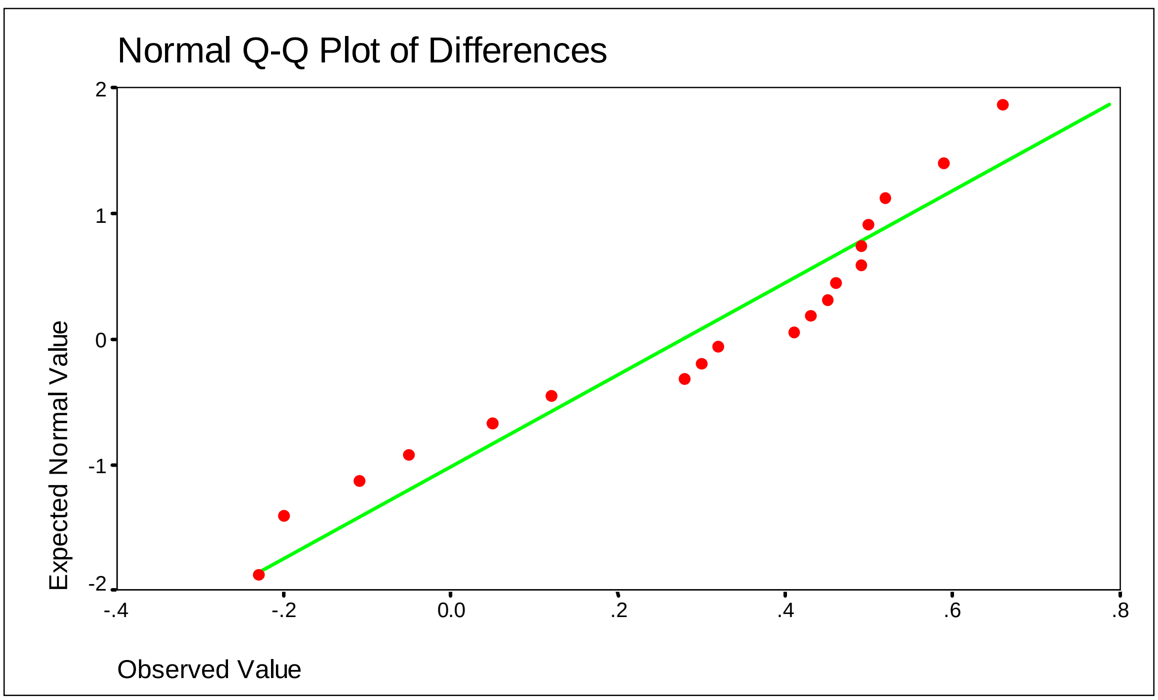
In other words, the null hypothesis states that there is no difference between the average tread depth of the brand A and brand B tires. The alternative hypothesis states that the average tread depth has increased and therefore the durability has improved.

The test statistic is a one-sample t, since we are now analyzing a single sample of differences:

$$t = \frac{\overline{D} - \mu_D}{s_D / \sqrt{n}}, \text{ where } df = n-1.$$

Here $\overline{D}$ is the sample mean of the differences, μ_D is the hypothesized value of the population mean, s_D is the sample mean of the differences, and n is the sample size.

Before we will apply the test, we have to check whether the assumption of normality is not violated. The normal quantile plot displayed below indicates that the assumption might be slightly violated.



Though the t-tools are quite robust against nonnormality, they should be used with some caution in the experiment.

The Paired-Samples T Test in SPSS produces the following output:

PAIRED SAMPLES STATISTICS				
VARIABLES	NUMBER	MEAN	STD. DEVIATION	STD. ERROR MEAN
BRAND A	20	6.4370	0.622	0.139
BRAND B	20	6.1605	0.616	0.138

VARIABLE	PAIRED DIFFERENCES			INFERENCES				
	Mean	Std. Deviation	Std. Error Mean	95% CI		TEST		
				Lower	Upper	t	Df	Sign
D= A-B	.2765	.2734	.0611	.1486	.4044	4.52	19	.000

The value of the test statistic t=4.52 expresses the distance between the hypothesized value of the population mean μ_D=0 and the observed sample mean $\overline{D}$ in standard errors. The p-value of the two-sided test is zero, thus the p-value of the one-sided test considered here is zero, too. Thus the data provide strong evidence that the durability has increased after the new technology was implemented. However, we cannot claim that the better durability is due to the new technology. The 95% confidence interval for the change in the tread depth is (0.1486, 0.4044) in 1/32 of an inch.

REMARK:

In our analysis above, we have made inferences about a single population, the population of all differences within matched pairs. It is inappropriate to use the two-sample t-test because the assumption of independent samples is invalid. The experiment was carried out in the way that makes the observations within pairs dependent.

14.5 Using the Nonparametric Methods

The paired t-test presented in the previous section has an assumption of normality for the differences. The normal quantile plot discussed in Section 14.4 indicates that the assumption of normality might be slightly violated for these data. Now we will use the Wilcoxon signed-rank test and the Sign test to make inferences

The Wilcoxon signed-rank test does not require normality of the differences, but it is assumed that the differences are independent and symmetric. The Wilcoxon signed-rank test can be used with normal and nonnormal data and is often more powerful than the paired t-test when the population is not normal. It is worthy to remind you that the power of a test is the probability of rejecting a false null hypothesis. A good test has high power.

SPSS produces the following output:

Wilcoxon Matched-Pairs Signed-Ranks Test

Mean Rank	Sum of Ranks	Cases
11.91	190.5	16 - Ranks (BRAND_B LT BRAND_A)
4.88	19.50	4 + Ranks (BRAND_B GT BRAND_A)
		0 0 Ties (BRAND_B EQ BRAND_A)

Number of observations 20

Z = -3.1925 2-Tailed P = .0014

The p-value of the two-tailed test is 0.0014. Thus the p-value of our one-sided test is 0.0014/2=0.0007. That small p-value indicates strong evidence against the null hypothesis in favour of the alternative μ_A-μ_B>0. However, observe that the assumption of symmetry for the distribution of differences is questionable. In our sample, 16 out of 20 differences are positive and only 4 are negative.

The Sign-Test counts the number X of pairs with a positive difference. The following output can be obtained with SPSS:

Sign Test

BRAND_A with BRAND_B

Cases

Cases		
16	- Diffs (BRAND_B LT BRAND_A)	
4	+ Diffs (BRAND_B GT BRAND_A)	
0	Ties	
--		(Binomial)
20	Total	Exact 2-Tailed P = .0118

The p-value for the one-sided alternative can be obtained by dividing the p-value for the two-sided alternative by two: 0.0118/2=0.0059. That small p-value strongly provides strong evidence against the null hypothesis of no difference between the durability of the two brands of tires.

14.6 Summary

In order to isolate and measure the effects of tire durability on tire wear, we used matched pairs design and restated the data as single-sample data by taking the differences within each pair. We applied the statistical tools to the differences.

The scatterplot of tread depths of brand A versus brand B tires, the boxplot and histogram of differences strongly support the thesis that the tread depth of brand A tires tends to exceed the tread depth of brand B tires.

The normal quantile plot for the differences indicates that the assumption of normality necessary to apply the t-tools might be slightly violated. Though the t-tools are quite robust against nonnormality, they should be used with some caution in the experiment. The p-value of the t-test is zero providing strong evidence that brand A tires have better durability than brand B tires, on the average. Thus the data provide strong evidence that the durability has increased after the new technology was implemented. The mean difference between the tread depth of the two brands is estimated to be between 0.1486 and 0.4044 in 1/32 of an inch. The conclusions are consistent with the results provided by distribution-free procedures, the Wilcoxon Signed-Rank Test and the Sign Test.

Can we make any conclusions about the effects of the new technology on tire wear? Can we claim that the new technology has improved the durability of tires? The answer to the above question depends on the conditions the new technology was implemented under. The new technology might be implemented in different conditions than the old technology, in a different factory building, by a different group of workers, and so on. It is possible that the smaller tire wear could be attributed to these factors, but not to the new technology.