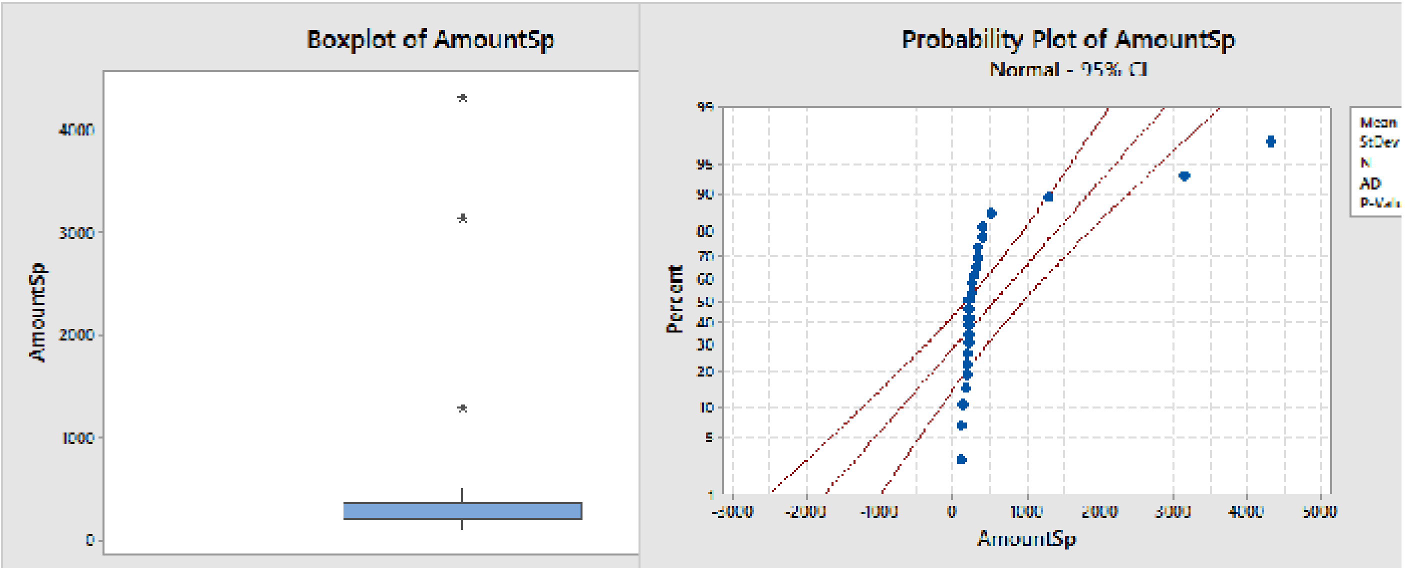
5.51

a. The normal probability plot and boxplot are given here:



The data set does not appear to be a sample from a normal distribution, since a large proportion of the values are outliers as depicted in the boxplot and several points are a considerable distance from the line in the normal probability plot. The data appears to be from an extremely right-skewed distribution.

b. Because of the skewness, the median would be a better choice than the mean.

Sign CI: AmountSp

Sign confidence interval for median

	N	Median	Achieved Confidence	Confidence Interval Lower	Confidence Interval Upper	Position
AmountSp	25	225.0	0.8922	211.0	326.0	9
			0.9500	208.6	338.8	NLI
			0.9567	208.0	342.0	8

```
> SIGN.test(exer5.21$AmountSpent)
```

c. (208, 342) We are 95% confident that the median amount spent on healthcare by the population of hourly workers is between \$208 and \$342.

Sign Test for Median: AmountSp

Sign test of median = 400.0 versus > 400.0

	N	Below	Equal	Above	P	Median
AmountSp	25	20	1	4	0.9999	225.0

d. Reject $H_0: M \le 400$ if $B \ge 25 - 7 = 18$.

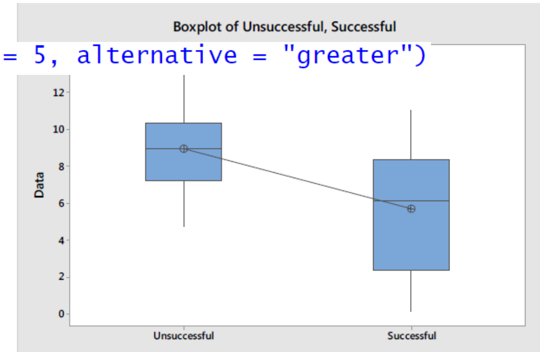
We obtain B = 4. Since 4 < 18, do not reject $H_0: M \le 400$. The data fails to demonstrate that the median amount spent on health care is greater than \$400.

```
> SIGN.test(exer5.21$AmountSpent,md = 400,alternative = "greater")
```

6.7

a. $H_0: \mu_U \geq \mu_S$ versus $H_a: \mu_U < \mu_S$; p-value < 0.0005 $\Rightarrow$ The data provide sufficient evidence to conclude that successful companies have a lower percentage of returns than unsuccessful companies.

b. The boxplots indicate that both data sets appear to be from normally distributed distributions; however, the successful data set indicates a higher variability than the unsuccessful.




```
> t.test(ex6.7$Unsuccess,ex6.7$Success, mu = 5, alternative = "greater")
```

c. The results from Minitab are below

```
Two-sample T for Unsuccessful vs Successful

              N  Mean  StDev  SE Mean
Unsuccessful  50  8.97   2.20     0.31
Successful    50  5.72   3.24     0.46


Difference = μ (Unsuccessful) - μ (Successful)
Estimate for difference:  3.251
95% lower bound for difference:  2.329
T-Test of difference = 5 (vs >): T-Value = -3.16  P-Value = 0.999  DF = 86
```

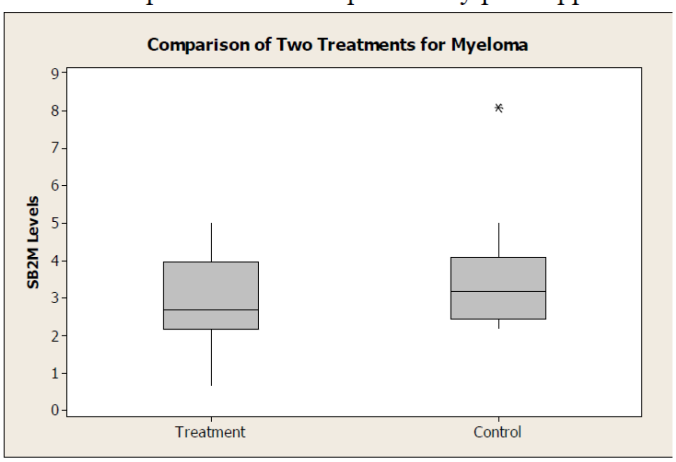
There is not significant evidence that the percentage for successful businesses returned goods is 5% less than that of unsuccessful businesses.

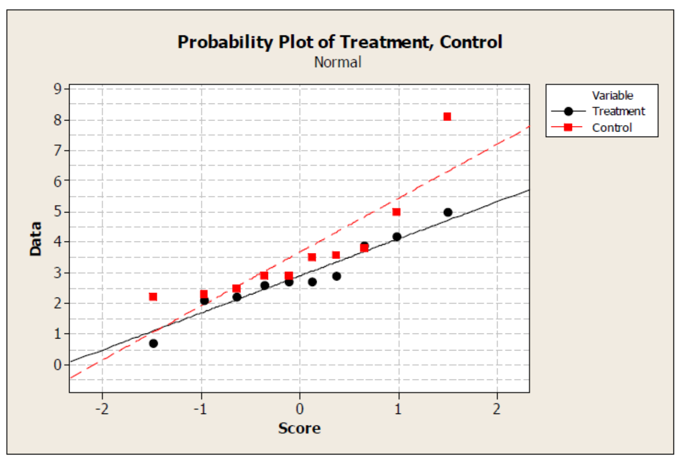
d. `95% CI for difference:  (2.149, 4.353)`

```
> t.test(ex6.7$Unsuccess,ex6.7$Success)
```

6.18

a. Boxplots and normal probability plots appear here:





Based on the plots, the treatment data appears to be from a normally distributed population, but the control data is right skewed with one large outlier.

b. The Wilcoxon rank sum statistic because of the small sample size and possible lack of normality.

c. H_0: The distributions are the same versus H_a: The distributions are different. Using the Wilcoxon rank sum statistic with $\alpha = 0.05$, reject H_0 if $T < 79$ or $T > 131$, where T is the sum of the ranks for the treatment group.

$T = 93 \Rightarrow$ fail to reject H_0. There is not significant evidence of a difference between the treatment and control groups.

d. The addition of alpha interferon (sumiferon) to the treatment regimen did not significantly change the effect on patient outcome.

```
> wilcox.test(ex6.18$Treatment, ex6.18$Control, alternative = "less")
```

6.29 `> t.test(ex6.29$Academic, ex6.29$NonAcademic, paired = T)`

a. $H_0: \mu_d = 0$ versus $H_a: \mu_d \neq 0$

$t = 4.95$, df = 29 $\Rightarrow$ p-value = $2P(t \geq 4.95) < 0.001$ There is significant evidence of a difference in the mean final grades.

b. A 95% confidence interval estimate of the mean difference in mean final grades is (2.23, 5.37).

c. We would need to verify that the differences in the grades between the 30 twins are independent. The normal probability plot would indicate that the differences are a random sample from a normal distribution. Thus, the conditions for using a paired *t*-test appear to be valid.

d. Yes. The purpose of pairing is to reduce the subject-to-subject variability, and there appears to be considerable differences in the students in the study. Also, a scatterplot of the data yields a strong positive correlation between the scores for the twins.

6.36 H_0: The distribution of differences (Benzedrine minus placebo) is symmetric about 0 versus H_a: The differences (Benzedrine minus placebo) tend to be larger than 0.

Wilcoxon Signed Rank Test: Difference

```
Test of median = 0.000000 versus median > 0.000000

                 N for   Wilcoxon          Estimated
             N    Test   Statistic      P     Median
Difference  14      14        16.0  0.012     +6.000
```

With $n = 14$, $\alpha = 0.05$, $T = T_-$, reject H_0 if $T_- \leq 25$.

From the data, we obtain $T_- = 16 < 25$ and thus reject H_0 and conclude that the distribution of heart rates for dogs receiving Benzedrine is shifted to the right of the dogs receiving the placebo. `> wilcox.test(ex6.36$Placebo, ex6.36$Benzedrine, paired = T, alternative = "less")`

6.37 One-sided research hypothesis: $\mu_T > \mu_P$; $\sigma \approx 18.6$; $\alpha = 0.05$; $\beta \leq 0.20$ whenever $\mu_T - \mu_P > 5$; $n_T = n_P = n$

$$n \approx \frac{2\sigma^2(z_\alpha + z_\beta)^2}{\Delta^2} = \frac{2(18.6)^2(1.645 + 0.84)^2}{5^2} = 170.9 \Rightarrow n = 171$$

6.40

a. $$n \approx \frac{\sigma_d^2(z_\alpha + z_\beta)^2}{\Delta^2} = \frac{(20)^2(2.326 + 1.28)^2}{10^2} = 52.01 \Rightarrow n = 53$$

b. We assumed that the distribution of the differences was normal.