Solutions

1 Since we have only 10 samples, we cannot apply the central limit theorem unless the samples come from a normal distribution. This is indeed the case here. We have to use the t-statistic, because the true variance is unknown.

a. The null hypothesis is $H_0\colon \mu = 50$. So we have

Test size:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{48 - 50}{\frac{4}{\sqrt{10}}} = -1.5811 \sim t(n-1)$$

Approximate p-value:

$$2 \cdot \left(1 - t_{cdf}(|t|, n-1)\right) = 2(1 - tcdf(1.5811, 10 - 1) = 2(1 - 0.9258) = 0.1483$$

And we fail to reject the null hypothesis.

b. Test size:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{48 - 50}{\frac{4}{\sqrt{10}}} = -1.5811 \sim t(n-1)$$

We have

```
t0 = tinv(1-0.05/2,n-1) = tinv(0.975,10-1) = 2.2622
```

The endpoints of the 95% confidence interval become

$$\mu_- = \bar{x} - t_0 \cdot \frac{s}{\sqrt{n}} = 48 - 2.2622 \frac{4}{\sqrt{10}} = 45.1385$$

$$\mu_+ = \bar{x} + t_0 \cdot \frac{s}{\sqrt{n}} = 48 + 2.2622 \frac{4}{\sqrt{10}} = 50.8615$$

2 The minimum required number of samples is

$$n \geq \left(\frac{1.96 \cdot \sigma}{B}\right)^2$$

Where $\sigma = 2.5$ and $B = 0.5$. Hence, we have

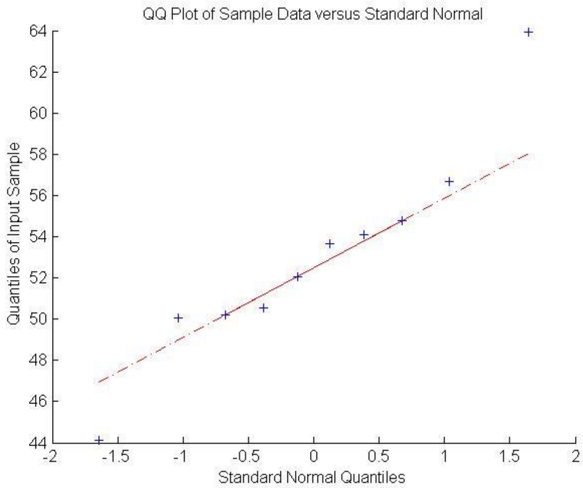
$$n \geq \left(\frac{1.96 \cdot \sigma}{B}\right)^2 = \left(\frac{1.96 \cdot 2.5}{0.5}\right)^2 = 96.0400$$

and conclude that we need at least 97 samples (because $n = 97$ is the smallest integer that is larger than or equal to 96.04).

3 This is an example, where *n* is small, so we have to make inference about the mean using the t-score, we have to check that the data are normally distributed.

a. Make a Q-Q plot

```
x = [ 54.0748
      56.6827
      54.7552
      44.1039
      50.2046
      53.6727
      63.9488
      50.5385
      50.0734
      52.0398]
qqplot(x)
```



The Q-Q plot results in a straight line, and hence we can conclude that the data are normally distributed.

b. The sample mean is

```
>> mean(x)
ans =
   53.0094
```

and the empirical variance is

```
>> var(x)
ans =
   26.7363
```

c. The empirical standard deviation is sqrt(26.7363) = 5.1707. We have n = 10 samples, and

```
t0 = tinv(1-0.05/2,n-1) = tinv(0.975,10-1) = 2.2622
```

The endpoints of the 95% confidence interval become

$$\mu_- = \bar{x} - t_0 \cdot \frac{s}{\sqrt{n}} = 53.0094 - 2.2622 \frac{5.1707}{\sqrt{10}} = 49.3104$$

$$\mu_+ = \bar{x} + t_0 \cdot \frac{s}{\sqrt{n}} = 53.0094 + 2.2622 \frac{5.1707}{\sqrt{10}} = 56.7084$$