Solutions

1 Let X be the random variable denoting the daily demand, at let $X \sim N(\mu, \sigma^2) = N(100, 15^2)$.

a. Standardizing, we get $\Pr(X > 125) = \Pr\left(Z > \frac{x-\mu}{\sigma}\right)$, where Z is standard normally distributed, $Z \sim N(0,1)$. Inserting, we get $\Pr\left(Z > \frac{x-\mu}{\sigma}\right) = \Pr\left(Z > \frac{125-100}{15}\right) = \Pr(Z > 1.6667) = 1 - \Pr(Z \leq 1.6667) = 1 - \Phi(1.6667) = 1 - 0.9522 = 0.0478$.

b. By symmetry of the normal distribution, the probability below 75 is the same as the probability above 125 (both are 25 away from the mean) so the answer is again 0.0478. For 70, standardize to get $\Pr(X \leq 70) = \Pr\left(Z \leq \frac{70-100}{15}\right) = \Pr(Z \leq -2) = \Phi(-2) = 1 - \Phi(2) = 1 - 0.9772 = 0.0228$.

c. The z-value for 0.95 is $z = \Phi^{-1}(0.95) = 1.6449$, so we need to set the inventory level 1.6449 standard deviations above the mean, at $\mu + z \cdot \sigma = 100 + 1.6449 \cdot 15 = 124.67$.

2 Let $\bar{X}$ denote the sample mean calculated from n=10 samples drawn from $X \sim N(\mu, \sigma^2) = N(50, 4^2)$.

a. According to the central limit theorem $\bar{X} \sim N(\mu, \sigma^2/n) = N(50, 4^2/10) = N(50,1.6)$. Hence, the expected value of $\bar{X}$ is 50, and the standard deviation is $\sqrt{1.6} = 1.2649$.

b. Matlab code:

```
mu      = 50;   % Population mean
sigma   = 4;    % Population standard deviation
n       = 10;   % Number of samples
x       = randn(1,n)*sigma + mu; % Random samples
x_bar   = mean(x)
s       = std(x)
```

c. Matlab code (continued):

```
for i = 1:100
    x       = randn(1,n)*sigma + mu; % Random samples
    x_bar(i)= mean(x);
end

mean(x_bar)
ans =
   50.1305

std(x_bar)
ans =
    1.24
```

We observe that the estimated mean (50.1305) is close to the theoretical mean (50), and that the estimated standard deviation (1.24) is fairly close to the theoretical standard deviation (1.2649).

3 Test statistic

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{48 - 45}{4/\sqrt{10}} = 2.3717 \sim N(0,1)$$

a. For a two-tailed test (H_1: $\mu \neq 45$), the p-value is

$$\begin{aligned} 2 \cdot \min\{\Pr(Z \geq z), \Pr(Z \leq z)\} \\ &= 2 \cdot \min\{2 - \Pr(Z \leq z), \Pr(Z \leq z)\} \\ &= 2 \cdot \min\{1 - \Phi(\mathrm{z}), \Phi(\mathrm{z})\} \\ &= 2 \cdot \min\{1 - \Phi(2.3717), \Phi(2.3717)\} = \\ &= 2 \cdot \min\{1 - 0.9911, 0.9911\} \\ &= 0.0177 \end{aligned}$$

The significance level is 5%, so p<0.05 and we reject the null hypothesis.

b. This is a right-tailed test (H_1: $\mu > 45$). The p-value is

$$\Pr(Z \geq z) = 1 - \Pr(Z \leq z) = 1 - \Phi(\mathrm{z}) = 1 - \Phi(2.3717) = 1 - 0.9911 = 0.0089$$

The significance level is 5%, so p<0.05 and we reject the null hypothesis.

4 For H_0: $\mu \geq 300{,}000$ versus H_1: $\mu < 300{,}000$ the test statistic is

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{350000 - 300000}{\frac{60000}{\sqrt{40}}} = 5.2705 \sim N(0,1)$$

This is a left-tailed test (H_1: $\mu < 300{,}000$). The p-value is

$$\Pr(Z \leq z) = \Phi(\mathrm{z}) = \Phi(5.2705) = 1$$

We fail to reject the null hypothesis, because p > 0.05. In words, it is extremely significant that the population mean is greater than 300,000.

5 The 95% confidence interval is

$$\bar{x} - 1.96 \cdot \sigma/\sqrt{n} \leq \mu \leq \bar{x} + 1.96 \cdot \sigma/\sqrt{n}$$

Inserting, we get

$$48 - 1.96 \cdot 4/\sqrt{10} \leq \mu \leq 48 + 1.96 \cdot 4/\sqrt{10}$$

and the confidence interval becomes

$$[\mu_-; \mu_+] = [45.5;\ 50.5]$$

6 The 95% confidence interval is

$$\bar{x} - 1.96 \cdot \sigma/\sqrt{n} \leq \mu \leq \bar{x} + 1.96 \cdot \sigma/\sqrt{n}$$

Inserting, we get

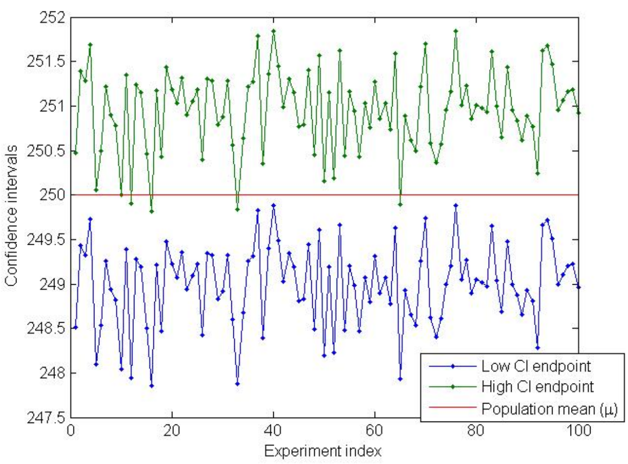
$$350000 - 1.96 \cdot 60000/\sqrt{40} \leq \mu \leq 350000 + 1.96 \cdot 60000/\sqrt{40}$$

and the confidence interval becomes

$$[\mu_-; \mu_+] = [331000;\ 369000]$$

7 Matlab code

```
n        = 25;    % number of cups
sigma    = 2.5    % standard deviation
mu       = 250;   % mean if machine is correctly calibrated
mu_obs   = 250.2; % observed sample mean
mu_hat   = [];    % simulated sample means

% Simulate 100 repeated experiments
for experiment = 1:100

    % 25 observations drawn randomly from the normal distribution
    X = randn(1,n)*sigma + mu;

    % Lower endpoint of 95% confidence interval
    mu_low(experiment) = mean(X) - 1.96*sigma/sqrt(n);

    % Upper endpoint of 95% confidence interval
    mu_high(experiment) = mean(X) + 1.96*sigma/sqrt(n);
end

% Plot the result of all 100 experiments
plot(1:100,mu_low,'.-',...
     1:100,mu_high,'.-',...
     [1 100],[mu mu])
legend('Lower CI endpoint','Upper CI endpoint','Population mean (\mu)')
xlabel('Experiment index')
ylabel('Confidence intervals')
```



In the example to the left, the true mean (red line) is included in the 95% confidence interval (between green and blue line) in 96 of the 100 experiments. That is, there are 4 four simulations, where the true mean does not lie within the confidence interval.

On average, the true mean will be included in 95 of the 100 confidence intervals. This is how we interpret the meaning of a 95% confidence level.