Statistics and Data Analysis
Suggested Solution for Homework 3

Instructor: Ling-Chieh Kung
Department of Information Management
National Taiwan University

1. The following R codes solve this problem:

```
x <- c(1, 2, 3, 4)
p <- c(0.5, 0.25, 0.125, 0.125)

barplot(p, names.arg = x, xlab = "x", ylab = "Probability") # (a)

mu.x <- sum(x * p) # (b)
var.x <- sum((x - mu.x)^2 * p) # (c)
sd.x <- sqrt(var.x) # (d)
```

   (a) The probability distribution is depicted in Figure 1.

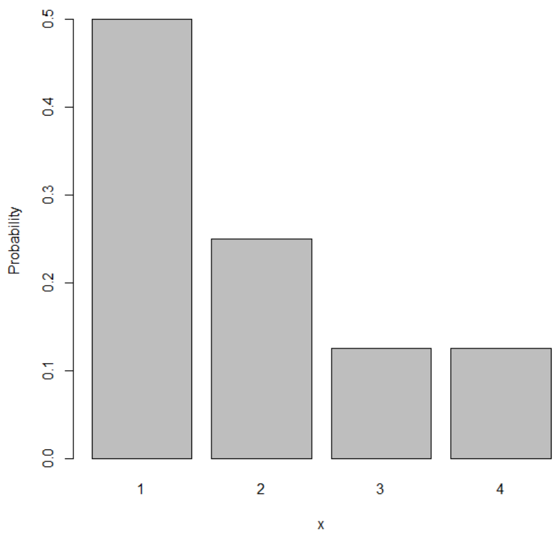

Figure 1: Probability distribution for Problem 1a

   (b) $\mathbb{E}[X] = 1.875$.
   (c) $\text{Var}(X) = 1.109$.
   (d) The standard deviation is $\sqrt{1.109} = 1.053$.

2. (a) $\Pr(X = 3) = 0$.
   (b) $\Pr(X \geq 20) = 0$.
   (c) $\Pr(X \geq 6) = 0.4$.
   (d) $\Pr(2 \leq X \leq 7) = 0.5$.
   (e) The pdf of X is $f(x) = 0.1$ for all $x \in [0, 10]$.

3. (a) The 10 integers between 1 and 10 are randomly drawn in order in a nonrepeating manner. In effect, the 10 integers between 1 and 10 are randomly permuted.
   (b) Now the 10 integers are randomly randomly drawn in a repeating manner. Some values may show up for multiple time while some others do not show up.

(c) `sample(1:42, 1)` and `sample(1:42, 3)`.

4. Yes, especially when `trial` is set to a large number.

5. Let X_1, X_2, and X_3 be the outcome of rolling three fair dices. What is the probability distribution of $Y = X_1 + X_2 + X_3$?

   (a) 3, 4, 5, ..., and 18.

   (b) Each time we conduct this experiment by running these codes, we see a figure similar to Figure 2. It is more likely to see values close to 11.

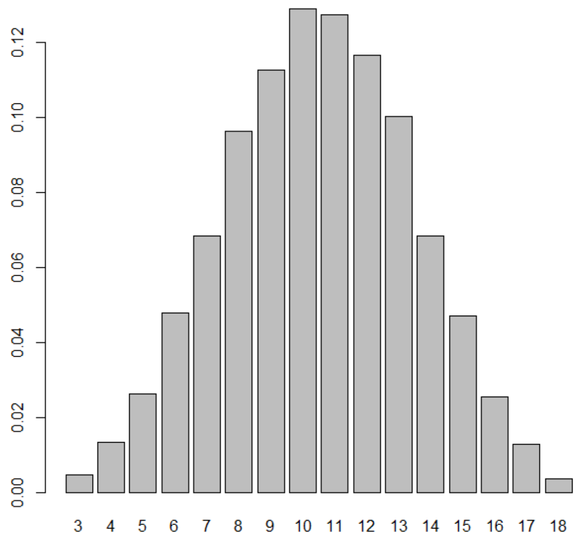

Figure 2: Probability distribution for Problem 5b

6. We have used the R function `pnorm(q, mean, sd)` to calculate the *left-tail probability* of normal distributions. Let's do some more practices.

   (a) $\Pr(X \leq 7) = 0.067$. This can be found by executing `pnorm(7, 10, 2)`.

   (b) $\Pr(X \geq 13) = 0.067$. This can be found by executing `1 - pnorm(13, 10, 2)`.

   (c) $\Pr(120 \leq X \leq 170) = 0.219$. This can be found by executing `pnorm(170, 200, 50) - pnorm(120, 200, 50)`.

   (d) $\Pr\left(\frac{120-200}{50} \leq Z \leq \frac{170-200}{50}\right) = 0.219$. In fact, we have

$$\Pr\left(\frac{a-200}{50} \leq Z \leq \frac{b-200}{50}\right) = \Pr(a \leq X \leq b)$$

   for any a and b.

7. Besides `pnorm()`, there are other functions related to the normal distribution:

   (a) To make the output look more like a normal distribution, we may increase the value of the first argument. For example, using 10000 will make the output quite like a normal distribution.

   (b) We will see that `p` is 0.05.

   (c) $\Pr(X \leq x) = 0.8$ when $x = 88.416$. This can be found by executing `qnorm(0.8, 80, 10)`.

8. The daily demand of a product $X \sim \text{ND}(80, 10)$. At the end of each day, you place an order to order q units from your supplier. The products will be ready at your store the next morning. Unsold products will become valueless.

(a) The probability is 0.067. This can be found by executing `1 - pnorm(95, 80, 10)`.

(b) Find the minimum integer q that achieves a 90% service level is 93. To see this, note that `qnorm(0.9, 80, 10)` returns 92.816. Therefore, the smallest integer that satisfies our requirement is 93.

(c) The correspondences between order quantities and service levels are illustrated in

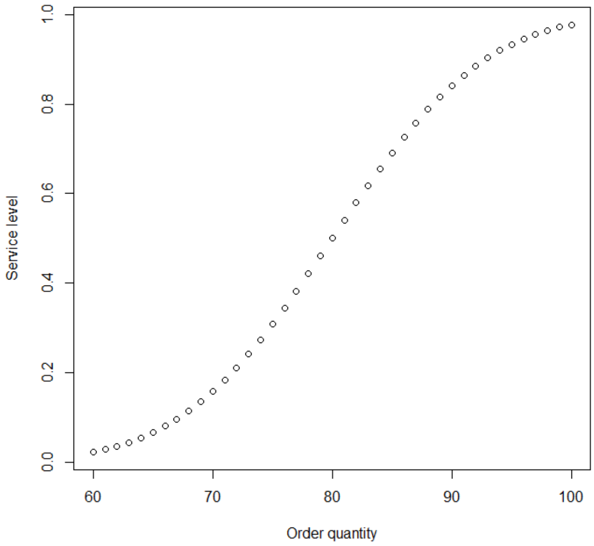

Figure 3: Probability distribution for Problem 8c

This figure can be generated by the following R codes:

```
q <- 60:100
s <- pnorm(q, 80, 10)
plot(x = q, y = s, xlab = "Order quantity", ylab = "Service level")
```

(d) The R code that solves Part (b) is `q[min(which(s >= 0.9))]`, where `q` and `s` have been defined in Part (c).