

Name: _____

Basic Statistics and Probability (2018-19)
Science & Engineering Program Boston University-Faculty of Science UAM

QUIZ 1

The total maximum score in this quiz is 10 points. The weight of this quiz in the final grade is 15%.

1. To determine the proportion of people in a certain town who are smokers (reported smoking at least 100 cigarettes in their lifetime and currently smoke every or some days), it has been decided to poll people at one of the following local spots:

- i) the pool hall;
- ii) the bowling alley;
- iii) the shopping mall;
- iv) the library.

The interviewed people will also be asked their age and the question

“In general, how harmful do you feel secondhand smoke is to adults – very harmful, somewhat harmful, not too harmful or not at all harmful?”

- a) (0.5 points)** Which of these potential polling places will most likely yield a representative sample? Why?
 - b) (0.25 points)** Identify the data collection method.
 - c) (0.5 points)** Identify the experimental units (individuals) and the population of the study.
 - d) (1 point)** Define the variables measured and classify them as qualitative or quantitative.
-

2. In a certain Spanish health care district, 12% of adolescent girls have iron deficiency. Suppose two adolescent girls are chosen at random in that district and their iron deficiency status is observed.

- a) (0.5 points)** List the sample points.
 - b) (0.5 points)** Compute the probability that one of the girls has iron deficiency and the other one does not.
 - c) (1 point)** Assign probabilities to the sample points.
-

3. In August 2010 there was a cave-in at the San José copper-gold mine in Chile. Thirty-three men were trapped 700 meters (2,300 ft) underground and 5 kilometers (3 mi) from the entrance of the mine. The ages (in years) of the miners were (sorted into ascending order):

```

19 23 24 27 27 27 29 29 30 31 31
33 33 34 34 40 40 43 44 44 45 46
48 48 49 50 52 53 54 54 56 56 63

```

In the following we display the R code and output used to analyze the data. You are allowed to use these outputs to solve the exercise. In the output of `summary(X)` we have substituted the resulting mean by question marks.

```

X = c(19, 23, 24, 27, 27, 27, 29, 29, 30, 31, 31,
      33, 33, 34, 34, 40, 40, 43, 44, 44, 45, 46,
      48, 48, 49, 50, 52, 53, 54, 54, 56, 56, 63)

```

```
summary(X)
```

```

  Min.   1st Qu.   Median   Mean   3rd Qu.   Max.
19.00   30.00   40.00   ????   49.00   63.00

```

```
var(X)
```

```
[1] 135.1098
```

```
sd(X)
```

```
[1] 11.62368
```

a) (0.25 points) The total sum of the observed ages is $\sum x = 1316$. Calculate the mean age of the trapped miners.

b) (0.75 points) What is the interquartile range of the data set? And the range? What is the percentage of observations between the two quartiles?

c) (1.5 points) Using the information given by `summary(X)`, compute the necessary elements to obtain the boxplot of the data. Draw it in the frame provided in Figure 1. Are there any outliers?

d) (1.25 points) Compute the absolute frequency of the following intervals:

[11,28], (28,38], (38,47], (47,55], (55,64].

Use your answer to draw a histogram of the sample in the space provided in Figure 2.

e) (1 point) What proportion of miners has an age between 11 and 28? What proportion of miners has an age greater than 38? Which are the absolute and relative frequencies of the interval [11,28]?

f) (0.25 points) Give the names and notation of the descriptive quantities computed by `var(X)` and `sd(X)`? Which is the relationship between these two quantities? **Indication:** It is not necessary to write down the mathematical formula to compute the quantities.

g) (0.25 points) Using the R outputs, compute the z -score corresponding to $x = 20$.

h) (0.5 points) What is the percentage of measurements in the data set that are above and below the 85th percentile? Is the 85th percentile of the sample of ages lower than 45? Justify your answer.

Indication: It is not necessary to compute the value of the 85th percentile.

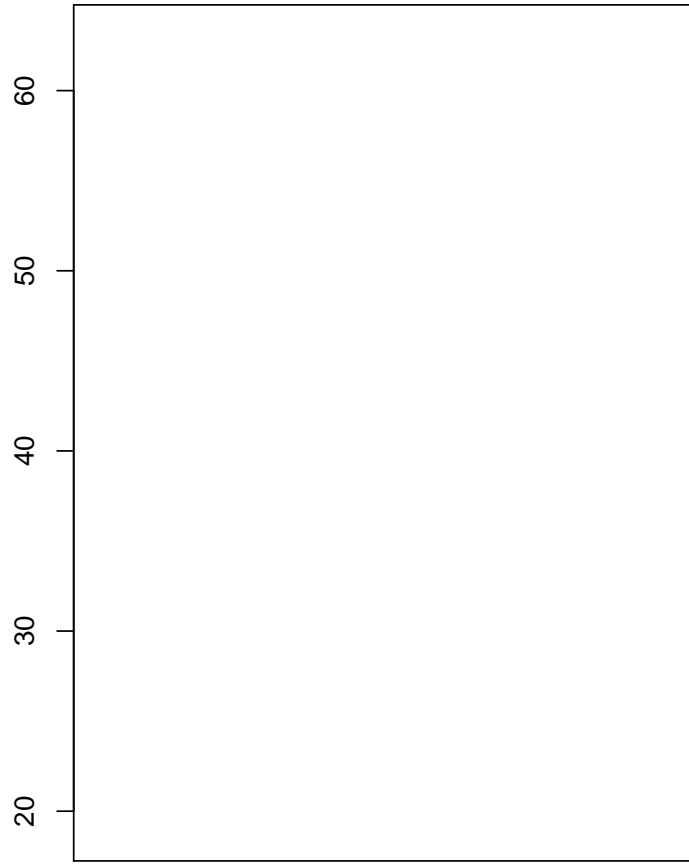


Figure 1

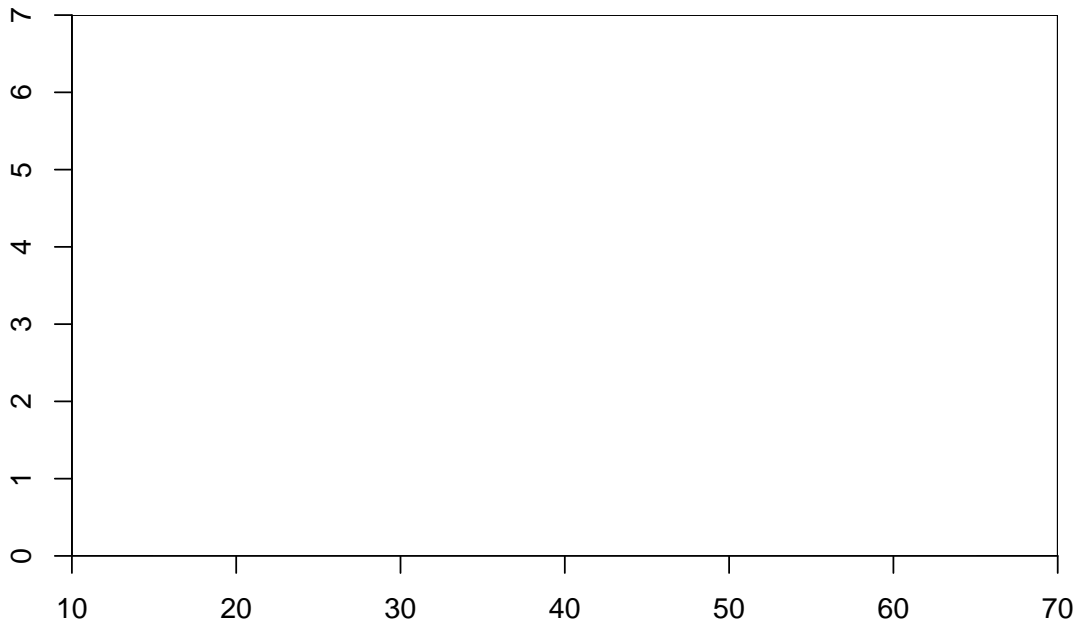


Figure 2