Solution Midterm test - 1 Stat2507 (D.E,F) Winter 2012

Multiple choice questions part: every question is worth 5 marks. Please circle only one answer..

The following three questions (1, 2, and 3) are based on the following stem-and-leaf plot for 31 observations.

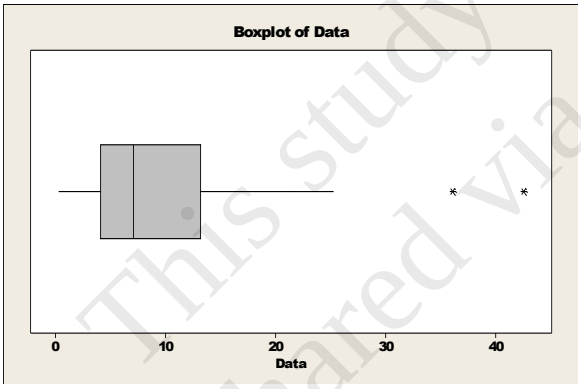
```
Stem-and-leaf of Data  N  = 31
Leaf Unit = 0.1

 3   6   036
 4   7   1
 7   8   399
 12  9   01335
 13  10  1
(7)  11  0156899
 11  12  8
 10  13  345
 7   14  0269
 3   15  09
 1   16
 1   17
 1   18  2
```

1. The lower (first) quartile and the median are
   a) 11.25 and 6 **b)** 9 and 6 **c)** 9 and 11.5 (**) **d)** 13.5 and 11

2. What is the approximate standard deviation of the data?
   a) 3.15 **b)** 3.05 (**) **c)** 9.3 **d)** need more information

3. The 7^{th} largest value is
   a) 8.9 **b)** 11.9 **c)** 14.0 (**) **d)** 14.6

The following two questions (4 and 5) are based on the following boxplot.



4. Which one of the following statements about the distribution above is correct:
   a) Symmetric and has no outlier **b)** Skewed to the left and has outliers(s) **c)** Skewed to the right and has outlier(s) (**) **d)** Symmetric but the mean and the median are different

5. Which one of the following statements is false?
   a) The mean is larger than the median **b)** The median is approximately equal 7
   c) The maximum value is approximately 43 **d)** The maximum value is approximately 25 (**)

6. Which of the following summary measures is affected by outliers?
   a) The first quartile **b)** The median **c)** The interquartile range (IQR)
   d) None of above (**)

7. The percentage of all observations in a data set that lie between the first and the third quartiles is
   a) 25 **b)** 50 (**) **c)** 75 **d)** 95.

8. Assume that 70% of Toronto residents believe that Toronto is a nice place to live. If we randomly interview 20 Toronto residents, what is the probability that more than 12 will say that Toronto is a nice place to live (use table attached)?
   a) 0.114 **b)** 0.886 **c)** 0.228 **d)** 0.772 (**)

9. If $P(A) = 0.8$, $P(B) = 0.7$, and $P(A \cup B) = 0.9$, then $P(A \cap B)$ is
   a.) 0.12 **b.)** 0.08 **c.)** 0.6 (**) **d.)** 0.5

10. If $P(A) = 0.7$, $P(B|A) = 0.5$, and $P(B) = 0.4$ then $P(A|B)$ is
    a.) 0.35 **b.)** 0.56 **c.)** 0.571 **d.)** 0.875 (**)

11. When sampling without replacement, the appropriate distribution is
    a) binomial distribution **b)** hypergeometric distribution (**) **c)** Poisson distribution,
    d) all of the above

12. Suppose that in a particular city, airport A handles 50% of all airline traffic, and airport B and C handle 30% and 20% respectively. The detection rates for weapons at the three airports are 0.9, 0.5, and 0.4 respectively. If a passenger is found to be carrying a weapon, what is the probability that he is using airport A?
    a) 0.85 **b)** 0.35 **c)** 0.66 (**) **d)** 0.72

Long answer questions. Show all your steps. Use blank spaces left for answers.

1. Consider the following set of observations: 2, 3, 1, 0, 4, 9.

   (a) **[12 marks]** Compute the mean $\bar{x}$, the variance s^2, and the standard deviation s for this set of data.
   Sol:
   $\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{19}{6} = 3.167$
   $s^2 = \dfrac{\sum x_i^2 - \frac{\sum x_i}{n}}{n-1} = \dfrac{111 - 19^2/6}{5} = 10.167$, and $s = \sqrt{s^2} = \sqrt{10.167} = 3.19$

   (b) **[4 marks]** Compute the z-score for the observation $x = 9$ and conclude as whether it is an outlier or not.

   Sol:
   $z = \dfrac{x - \bar{x}}{s} = \dfrac{9 - 3.167}{3.19} = 1.83$, so $x = 9$ is not an outlier.

2. Weight gain (rounded to the nearest 5 pounds) per month for a calf has the following distribution table

x	0	5	10	15
$p(x)$	0.1	?	0.3	0.1

q

(a) **[4 marks]** Find $p(5)$.
Sol:
$p(5) = 1 - (0.1 + 0.3 + 0.1) = 0.5$

(b) **[12 marks]** Compute the average weight gain per month μ and the standard deviation σ of the weight gain.

Sol:
$E(X) = \sum xp(x) = 0 + 2.5 + 3 + 1.5 = 7, \quad E(X^2) = \sum x^2 p(x) = 12.5 + 30 + 22.5 = 65$
$V(X) = E(X^2) - (E(X))^2 = 65 - 49 = 16, \ \sigma = \sqrt{16} = 4$

3. The number of received calls at a home phone has a Poisson distribution and averages 0.5 calls per hour.

(a) **[4 marks]** Find the probability of no calls during the next hour.
Sol: For Poisson random variable: $P(X = k) = \dfrac{e^{-\mu}\mu^k}{k!}$, and $\mu = 0.5$ per hour
$P(X = 0) = e^{-0.5} = 0.61$

(b) **[4 marks]** Find the probability of 2 calls during the next two hours.

Sol: $\mu = 1$ per two hour
$P(X = 2) = \dfrac{e^{-1}1^2}{2!} = 0.184$