STATISTICS

Past papers

Worked by: Anas Khraim

IDENTIFY COMPUTE INTERPRET

First Exam Past Papers

Chapter 1:

Q1:

In a distribution of 160 values with a mean of 72, at least 120 fall within the interval (67,77). Approximately the percentage of values that should fall in the interval (64.5, 79.5) is:

a) 75%
b) 93.75%
c) 88.88%
d) 96%
e) 55.55%

Ans: C

Q2:

Suppose each employee in the company receives a $3000 raise for the next year (each employee's salary is increased by $3000). The standard deviation of the salaries for the employees will:

a) be multiplied by $3000
b) increase by $3000
c) be unchanged
d) increase by $\sqrt{3000}$
e) be multiplied by $=3000^2$

Ans: C

Q3:

A set of data has the following five number summary:

Min	First quartile	Median	Third quartile	Max
17	37	40	49	90

Which of the following contains all the outliers in the distribution?

a) 83,85,90,95
b) 17,78,80,85,90
c) 64,80,85
d) 2,3,85,90
e) 0,80,84,89

Ans: B

Q4:

For the following grouped frequency distribution:

Class	0-5	6-11	12-17	18-23
Frequency	2	8	7	3

The number of observations that lie bellow 21.5 is

a) 19
b) 8
c) 6
d) 14
e) 10

Ans: A

Q5:

Which of the following are true statements?

I. The variance is the square root of the standard deviation.
II. The standard deviation is zero only when all values are the same.
III. The standard deviation is strongly affected by outliers.

a) I and II
b) I and III
c) I, II and III
d) II and III
e) None of the above gives the complete set of true responses.

Ans: D

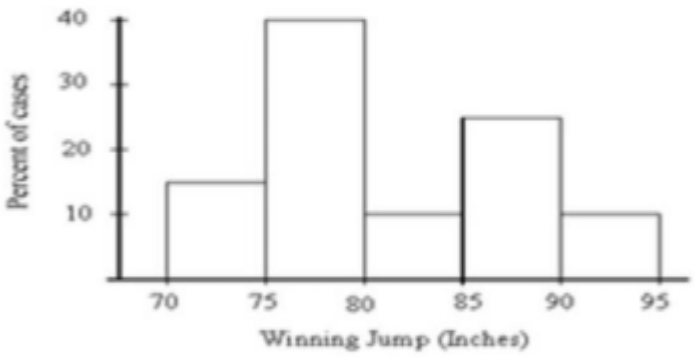
Q6:

If the heights of students has a symmetric distribution about the median Q2=170 cm. And the lower quartile Q1=155cm, then the iner-quartile range, IQR=

a) 30
b) 15
c) 185
d) 75
e) 50

Ans: A

Q7:



Above is the histogram of the gold medal winning high jumps for the Olympic Games. The 95th percentile of the histogram is approximately:

a) 75 inches
b) 81.25 inches
c) 92.5 inches
d) 82.5 inches
e) 85 inches

Ans: C

Q8:

If the frequency curve of a set data has a bell-shape with mean 200 and variance 100, then the percentage of observations that lies in the interval (170,190) is approximately equals to:

a) 81.5%
b) 83.5%
c) 2%
d) 13.5%
e) 15.5%

Ans: E

Q9:

A data set produced the five number summary shown below. Ther are no outliers in the data set.

Min	First Quartile	Median	Third Quartile	Max
22	31.2	44.5	59.8	67

Which of the following conclusions can be drawn from the data?

I. The mean is less than 44.5
II. Approximately 75% of the scores are bellow 59.8
III. Approximately 50% of the scores lie between 31.2 and 59.8

a) I only
b) II only
c) III only
d) I and III only
e) II and III only

Ans: E

Q10:

there are three children in a room, ages five, six and seven. If the six-year-old child enters the room, then:

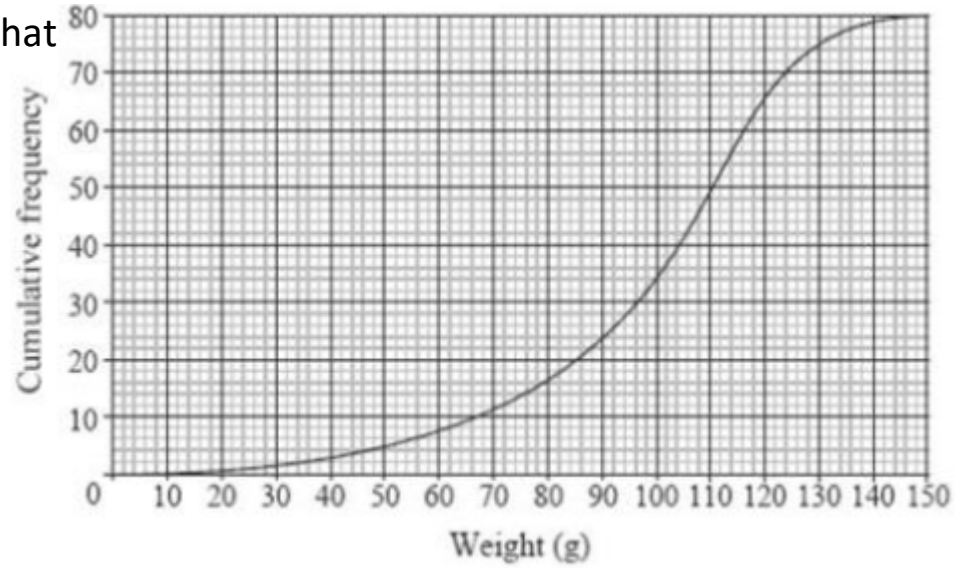
a) The mean age will stay the same but the variance will increase.
b) The mean age and the variance will stay the same.
c) The mean age will stay the same but the variance will decrease.
d) The mean age and variance will increase.
e) The mean age and variance will decrease.

Ans: C

Q11:

The cumulative frequency graph below represents the weight in grams of 80 apples picked from a particular tree

The number of apples that weigh more than 110 grams is:



a) 30
b) 50
c) 40
d) 85
e) 115

Ans: A

Q12:

From January to September, the mean number of car accidents per month was 630. From October to December, the mean was 810 accidents per month.

The mean number of car accidents per month the whole year was:

a) 720
b) 675
c) 690
d) 810
e) 630

Ans: B

Q13:

In a factory producing glasses, the weights of glasses are known to have a mean of 160 grams. It is also known that the interquartile range of the weights of glasses is 26 grams. Assuming the weights of glasses to be normally distributed, then the standard deviation of the weights of glasses equals:

a) 17.9
b) 26.9
c) 24.9
d) 20.9
e) 19.4

Ans: E

Chapter 2:

Q1:

the probability distribution of a discrete random variable X is defined by

$$P(X = x) = c\,x\,(4 - x), \quad x = 1, 2, 3, 4.$$

The value of c =

a) 0.1
b) 5
c) 0.5
d) 0.01
e) 0.05

Ans: A

Q2:

in a class of 20 students, 12 study Biology, 15 study History and 2 students study neither Biology nor History.

Given that a randomly selected student studies Biology, the probability that this student also studies History equals:

a) 0.25
b) 0.75
c) 0.60
d) 0.10
e) 0.80

Ans: B

Q3:

The random variable Y is such that $E(2Y + 3) = 6$ and $Var(2 - 3Y) = 11$.

$E(Y^2) =$

a) 47
b) $\frac{125}{36}$
c) 15
d) 147
e) $\frac{1213}{324}$

Ans: B

Q4:

Let A and B be 2 events such that P(A) = 0.2 and P(B) =0.4 and P(A U B) = 0.5. then the value of P($B \mid A$) =

a) $\frac{1}{3}$
b) 0.5
c) $\frac{1}{6}$
d) $\frac{2}{3}$
e) 0.25

Ans: B

Q5:

A corona virus is spreading through a city. A vaccination is available to protect against the virus. If a person has had the vaccination, the probability of catching the virus is 0.1; without the vaccination, the probability is 0.3. The probability of randomly selected person catching the virus is 0.24.

The probability that a randomly chosen person has been vaccinated is:

a) 0.55
b) 0
c) 0.40
d) 0.30
e) 0.045

Ans: D

Q6:

A bag contains 5 red and *n* green balls.

Two balls are chosen without replacement from this bag. If the probability that two red balls are chosen is $\frac{5}{18}$, then *n* =

a) 6
b) 5
c) 4
d) 13
e) 2

Ans: C

Q7:

A biased coin is weighted such that the probability of obtaining a head is $\frac{4}{7}$. The coin is tossed 6 times and ***X*** denotes the number of heads observed. The value of the ratio $\frac{P(X=3)}{P(X=2)}$=

a) 16
b) 9
c) $\frac{\mathbf{16}}{\mathbf{9}}$
d) $\frac{\mathbf{9}}{\mathbf{16}}$
e) 1

Ans: C

Q8:

Jenny goes to school by bus every day. When it is not raining, the probability that the bus is late is $\frac{3}{20}$. when it is raining, the probability that the bus is late is $\frac{7}{20}$. the probability that it rains on particular day is $\frac{9}{20}$. On one particular day the bus was late. The probability that it is raining on that day is:

a) $\frac{21}{32}$
b) $\frac{19}{20}$
c) $\frac{11}{20}$
d) $\frac{1}{20}$
e) $\frac{11}{32}$

Ans: E

Solutions Chapter 1

Q1

$n = 160 , \bar{x} = 72$

يتم حساب النسبة المئوية للبيانات المتواجدة داخل الفترة (67,77) كالتالي

$$\frac{120}{160} = \frac{3}{4} = 75\%$$

وبعلمنا بأن النسبة المئوية أساسا تُحسب من القانون $1 - \frac{1}{k^2}$ يمكن حساب قيمة ال k

$$1 - \frac{1}{k^2} = \frac{3}{4} \rightarrow \frac{k^2 - 1}{k^2} = \frac{3}{4} \rightarrow 4k^2 - 3k^2 = 4 \rightarrow \boldsymbol{k = 2}$$

وبما أن المتوسط يقع داخل الفترة (67,77) فتكون الأخيرة على الصورة

$$(\bar{x} - ks , \bar{x} + ks) \longrightarrow (72 - 2s , 72 + 2s)$$

$$72 - 2s = 67 \longrightarrow s = 2.5 , 72 + s = 77 , s = 2.5$$

و على نفس النمط الفترة المطلوب نسبة البيانات الموجود بداخلها على الصورة

$$(\bar{x} - ks , \bar{x} + ks) = (72 - 2.5k , 72 + 2.5k) = (64.5 , 79.5)$$

$$72 - 2.5k = 64.5 \longrightarrow k = 3 , 72 + 2.5k = 79.5 \longrightarrow k = 3$$

وبعلمنا بأن النسبة المئوية أساسا تُحسب من القانون $1 - \frac{1}{k^2}$

$$1 - \frac{1}{9} = 0.888889 \approx 88.88\%$$

Q2:

The increasing factor of data here isn't constant therefore, the standard deviation will remain the same as its only affected by multiplication and division.

Q3:

$$IQR = Q3 - Q1 = 49 - 37 = 12$$

The upper and lower for outlier's boundary is calculated through $(Q1 - 1.5IQR , Q3 + 1.5IQR) = (37 - 18, 49 + 18)$

$$= (19 , 67)$$

كل خارج عن الفترة فهو outlier ومع مراعاة أن الحد الأدنى للبيانات 17 والأقصى 90 يمكن استبعاد a لوجود 95 التي هي تكبر الحد الأقصى و c لعدم وجود outlier فيها أساسا و e و d لنفس سبب a

Q4:

من صياغة السؤال فالعدد 21.5 هو الpercentile لنسبة معينة فسنقوم بالعملية العكسية...وملاحظة أن 21.5 تقع بين الفترتين الأخيرتين

$$C.f_1 \quad \frac{kn}{100} \quad C.f_2 = 17 \quad \frac{20k}{100} \quad 20$$

$$Urb_1 \quad P_k \quad Urb_2 = 17.5 \quad 21.5 \quad 23.5$$

$$\frac{20-17}{\frac{k}{5}-17} = \frac{23.5-17.5}{21.5-17.5} \Rightarrow k = 95$$

أي أن 21.5 هو الP95 الذي يعبر عن العدد 95% من الأعداد وتحته 95% من البيانات

$$\frac{95}{100} * 20 = 19$$

Q5:

The first statement is **false** as the standard deviation IS equal to the square root of the variance not the other way around

Let's experiment with the second statement and assume $s = 0$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = 0 \Rightarrow x_i - \bar{x}\ MUST = 0\ , x_i = \bar{x}$$

Thus for every
x_i *it must be equal to the mean therefore identical*

For the third statement, to put it simply since the existence of outliers mean skewing the distribution curve leading to ultimately affecting the mean value (Skewing to the left or right can be used as proof for my explanation), and since we use the mean to calculate standard deviation, the latter, must be affected!

Thus, the second and third statements are true.

Q6:

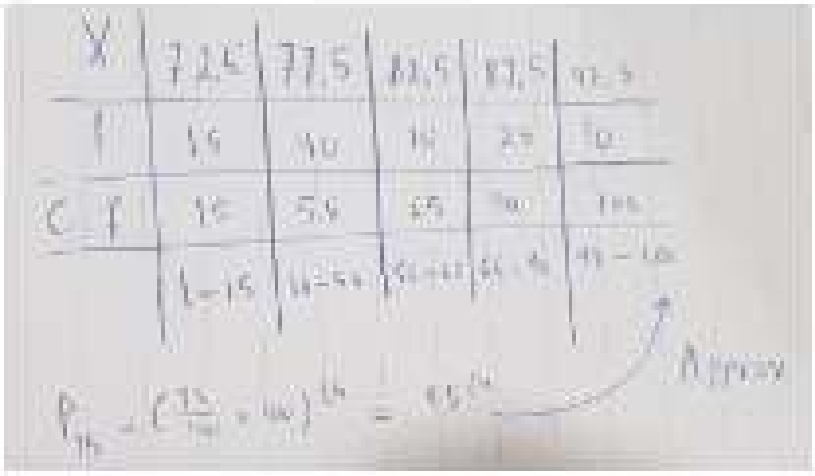
Because of the given symmetric distribution...

$$Q3 - Q2 = Q2 - Q1 \rightarrow 170 - 155 = 155 - Q1 \therefore Q1 = 140$$

$$IQR = Q3 - Q1 = 170 - 140 = 30$$

Q7:

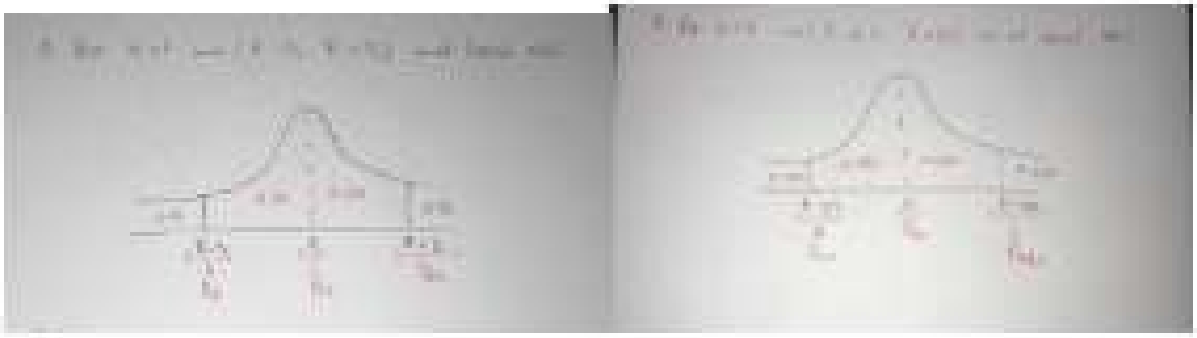
نحسب المتوسط لكل فترة ونصنع جدول ثاني كالتالي مع معاملة المحور الصادي على إنها frequency عادية



Q8:

$$s = \sqrt{s^2} = \sqrt{100} = 10$$

بما إن الفترة (170 , 190)المتوسط يقع خارجها ونلاحظ أنها على الصورة $(\bar{x} - 3s\ ,\bar{x} - s)$ فسنلجأ لـ Empirical rule ومن خلال المنحنيات التالية



فإن الفترة المحصورة بين $\bar{x}$-s و $\bar{x}$-3s هي ناتج طرح نسبهما الموضحة

$$P_{16} - P_{0.5} = 0.16 - 0.05 = 15.5\%$$

Q9:

To determine the distribution shape we must compare The difference between Q3,Q2 and Q2,Q1

$$Q3 - Q2 = 59.8 - 44.5 = 15.3$$

$$Q2 - Q1 = 44.5 - 31.2 = 13.3$$

$$\because Q3 - Q2 > Q2 - Q1$$

$$\therefore \textit{The distirbution is skewed to the right}$$

Thus Mode < Q2 < Mean , therefore $\bar{x} > 44.5$, thus, first statement is false

Because of 59.8 being Q3 second statement is true

Q3,Q1 represent the 75% , 25% of the data respectively and since IQR is the difference between them, thus an interval (Q3,Q1) would have 50% of data laying in it.

Q10:

أولا/نحسب متوسط الأعمار والVariance قبل دخول الطفل الجديد

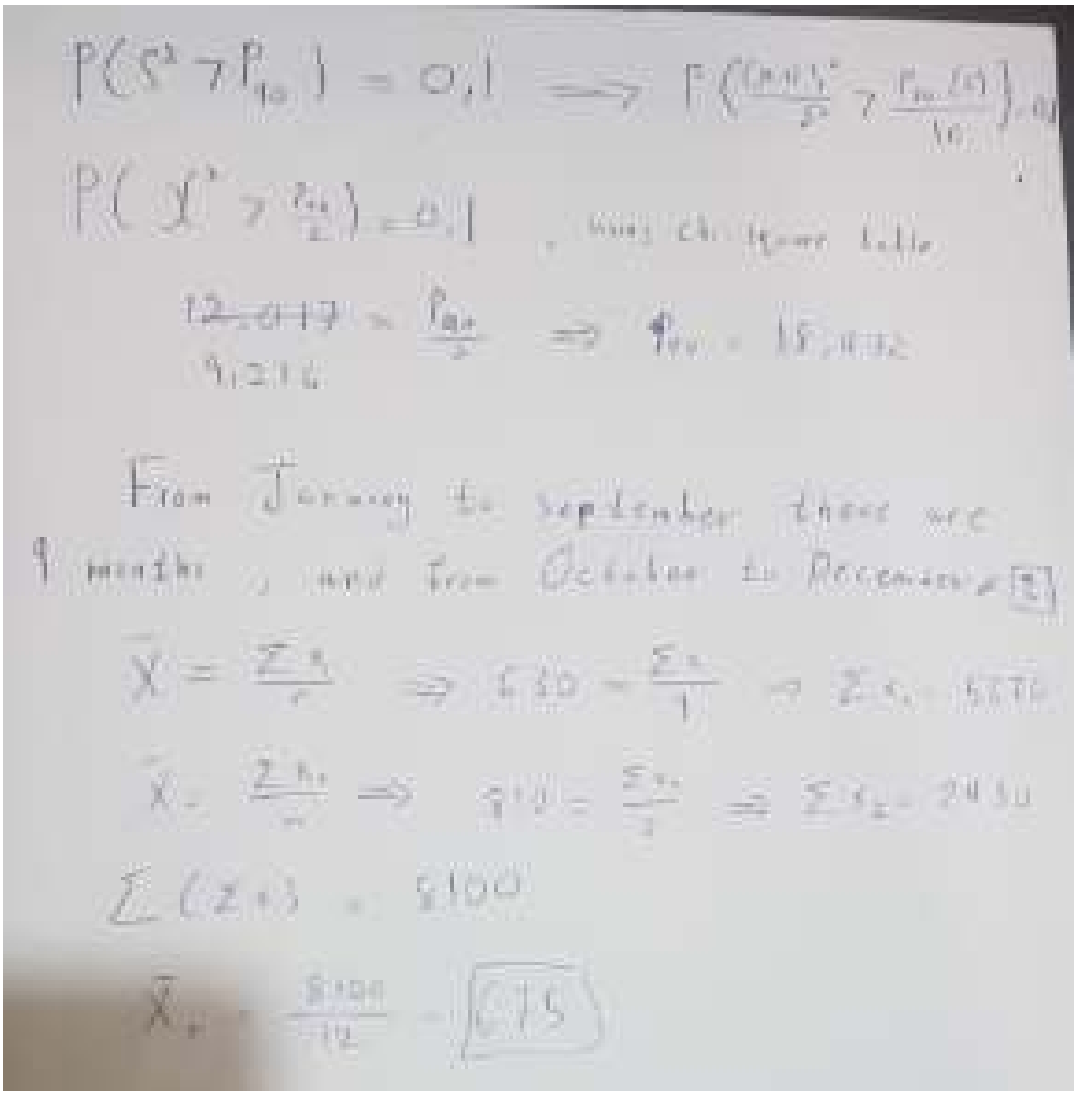
$$\bar{x} = \frac{5+6+7}{3} = 6 \ , S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)} = \frac{110}{2} - \frac{324}{6} = 1$$

أولا/نحسب متوسط الأعمار والVariance بعد دخول الطفل الجديد

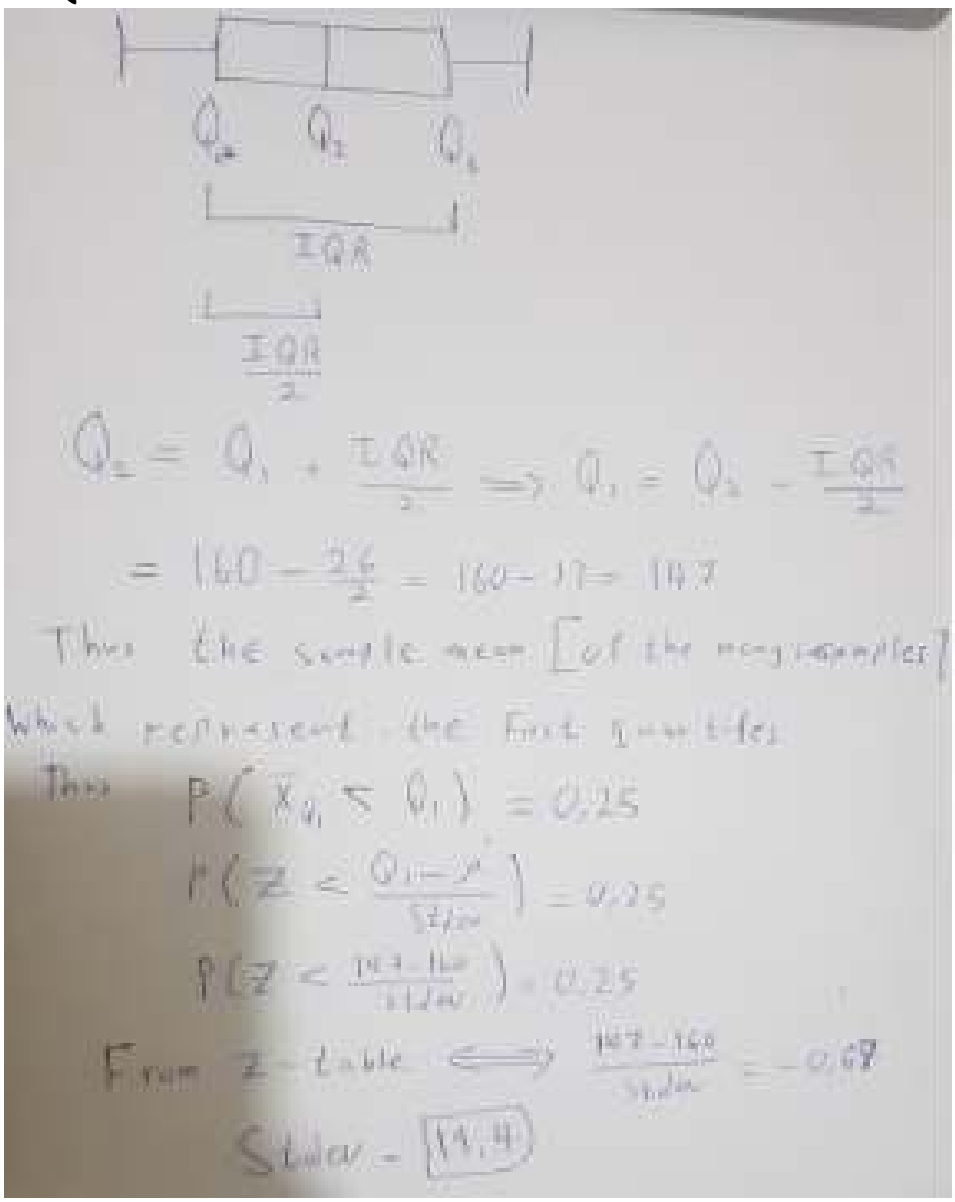
$$\bar{x}_1 = \frac{5+12+7}{4} = 6 \ , S_1^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)} = \frac{146}{3} - \frac{576}{6} = \frac{2}{3}$$

$$Since \ \bar{x} = \bar{x}_1 \ , S^2 > S_1^2$$

Q12:


From January to September there are
9 months , and from October to December [3]

Q13:


Q_1
Q_2
Q_3
IQR
$\frac{IQR}{2}$
$Q_2 = Q_1 + \frac{IQR}{2} \Rightarrow Q_1 = Q_2 - \frac{IQR}{2}$
$= 160 - \frac{26}{2} = 160 - 13 = 147$
Thus the sample mean [of the new samples]
which represent the first quartiles
Thus $P(\bar{X}_{Q_1} \leq Q_1) = 0.25$
$P(Z < \frac{Q_1 - \mu}{\sigma/\sqrt{n}}) = 0.25$
$P(Z < \frac{147 - 160}{\sigma/\sqrt{n}}) = 0.25$
From Z-table $\Longleftrightarrow \frac{147 - 160}{\sigma/\sqrt{n}} = -0.67$
$\sigma/\sqrt{n} = 19.4$

Solutions Chapter 2:

Q7:

$$\frac{P(X=3)}{P(X=2)} = \frac{\binom{6}{3} * \left(\frac{4}{7}\right)^3 * \left(\frac{3}{7}\right)^3}{\binom{6}{2} * \left(\frac{4}{7}\right)^2 * \left(\frac{3}{7}\right)^4} = \boxed{\frac{16}{9}}$$

Q8:

$P(L \mid NR) = \frac{3}{20}$, $P(L \mid R) = \frac{7}{20}$

$P(R) = \frac{1}{20} \Longrightarrow P(NR) = \frac{11}{20}$

$P(NR \mid L) = \frac{P(NR \cap L)}{P(L)}$

$P(L \mid NR) = \frac{P(NR \cap L)}{P(NR)} \Longrightarrow P(NR \cap L) = P(NR) \cdot P($

$P(NR \cap L) = \frac{11}{20} * \frac{3}{20} = \boxed{0,0525}$

$P(L \mid R) = \frac{P(L \cap R)}{P(R)} \Longrightarrow P(L \cap R) = 0,1575$

$P(L) = P(L \cap R) + P(L \cap NR) = 0,0525 + 0,1575$

$P(L) = 0,24$

$P(NR \mid L) = \frac{0,0525}{0,24} = \boxed{\frac{11}{32}}$