لجان النُّفعات

STATISTICS

MORPHINE ACADEMY

MORPHINE ACADEMY

Pharmaceutical Statistics

Lecture 6 Part 1

Descriptive statistics

Measures of Dispersion

Prepared and Presented by
Dr. Muna Oqal

Measures of Dispersion for Grouped Frequency Table

A. Range:

The range of grouped frequency table can be calculated as follow:

اذا كان في جدول

R=(Upper limit of last class – Lower limit for the first Class)+1

Measures of Dispersion for Grouped Frequency Table

- **Example**

The following frequency table represents the time in minutes of 50 pharmacists to their work selected from the records of a given pharmaceutical industries company in Jordan:

Time to travel to work	Frequency
1-10	8
11-20	14
21-30	12
31-40	9
41-50	7

لو كان قيم عشوائية
يعني بدون جدول:
القانون بكون بدون
زيادة واحد

A. Calculate the range for the time travelled to the work for pharmacists?

Solution:

R=(Upper limit of last class – Lower limit for the first Class)+1

R=(50-1)+1=50 minutes

Pharmaceutical Statistics

Lecture 6 Part 2

Descriptive statistics

Measures of Position

Prepared and Presented by
Dr. Muna Oqal

Measures of Position

Definitions

مقايس الموضع تستخدم لوصف

- **Measures of position are used to describe the relative location of an observation.**

موقع النسبي لملاحظة

نظام كميات موقع القيم

- **Quantiles system: are the positions of values in the data set. Data set are divided to quartiles, percentiles, percent (25% each), etc., after ordering them from smallest to the largest.**

تقسيم ارباع

نسب مئوية

Measures of Position

وسيط امثلة محددة

- The median and quartiles are specific examples of quantiles.

انظمة ربع اكثر من اربع

- Quantile systems that cut data into more than four ranges are really only useful where there are quite large numbers of observations. Such as quintiles, deciles and centiles (percentiles).

نطاقات مفيدة فقط عدد كبير من الملاحظات

- There are four quintiles, which divide data into five ranges, nine deciles for ten ranges and 99 centiles that produce 100 ranges.
- The ninth decile is thus equivalent to the 90th centile and both indicate a point that ranks 10% from the top of a set of values.

Measures of Position

انظمة الكميات

Quantile systems

مجموعات بيانات

Quantile systems divide ranked data sets into groups with equal numbers of observations in each group. Specifically: مجموعات تحتوي اعداد متساوية من ملاحظاتفي كل مجموعة

رباعيات

- 3 *Quartiles* divide data into four equal-sized groups.

خمسيات

- 4 *Quintiles* divide it five ways. بضيف واحد

اعشار

- 9 *Deciles* divide it ten ways.

- 99 *Centiles* divide it 100 ways.

Measures of Position

A. Percentiles
B. Quartiles
C. Five Number Summary
D. Boxplot

Measures of Position

- **Percentiles or (centiles): are measures which divide the data set into 100 equal parts (P_1, P_2,…,P_{100})**
- **Percentiles: are measures indicating the value below a given percentage of observations in group of observation falls.**

  تشير

  قيمة تقل عن نسبة مئوية معينة في مجموعة من ملاحظات

- **For example, the 40th percentile (P_{40}) is the value (or score) below which 40% of the observations may be found.**

Measures of Position

Finding the Score Given a Percentile

$$L = \frac{K}{100} \bullet n$$

n total number of values in the data set عدد الكلي في data

k percentile being used نسبة مئوية

L locator that gives the *position* of a value

P_k k th percentile

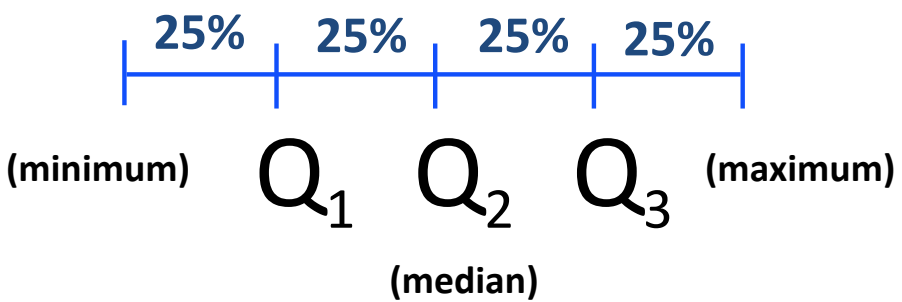
Then, you add 0.5 to find the exact position

اضيف ٥. عشان احدد موضع بزبط

Quartiles

Q_1, Q_2, Q_3

divides ranked scores into four equal parts



$Q_1 = P_{25}$ $Q_2 = P_{50}$ $Q_3 = P_{75}$

ربع النسبة مئوية

Q=quartile, P= Percentile

Measures of Position

A. The Quartiles for Raw Data (ungrouped data)

To find the quartiles for a set of raw data, do the following:

ارتب من اصغر للاكبر

1. Arrange the data set from the smallest to highest (ordered array).

احسب وسيط

2. Calculate the median (MD=Q2) for the data set.
3. For the half of the data set to the left of the MD calculate their median to get the first quartile (Q1).
4. For the half of the data set to the right of the MD, calculate their median to get the third quartile (Q3).

Measures of Position

- **Example (n is odd)**

The times in minutes needed for a random of 15 tablets to disintegrate are as follow:

(Q1) Median (Q2) (Q3)

5 10 10 10 10 12 15 20 20 25 30 30 40 40 60

Find the quartiles Q1, Q2 and Q3?

1. Data is already in an order from smallest to highest
2. The **MD** is $X_{((n+1)/2)} = X_{(8)} =$ **20** minutes.
3. For the half to the left of the median, n=8, then $X_{((n+1)/2)} = X_{(4.5)}$ = **Q1**=(10+10)/2=**10** minutes.
4. The second quartiles is **Q2**=**MD**=**20** minutes.
5. For the half to the right with the MD, n=8, then then $X_{((n+1)/2)} = X_{(4.5)}$ = the third quartile which is **Q3**=(30+30)/2=**30** minutes.
6. The quartiles are: **Q1=10 minutes**, **Q2=20 minutes**, **Q3=30 minutes**.

ملاحظة احنا ضمينا ٢٠ لل ربع الاول و ثاني لانو عدد فردي بعدين حسبنا median زي دايما

When n is odd, you can include the median in both quartiles **(Q_1 and Q_3)**

Measures of Position

Conclusion ٧٥٪ بتحتاج اكثر من ١٠دقايق

- This means that 25% of the tablets need less than 10 minutes to disintegrate. تفكك

- 50% of the tablets need 20 minutes to disintegrate.

قبل ٣٠ دقيقة يتم تفكيك ٧٥٪ من جميع اقراص

- Before 30 minutes 75% of all tables were disintegrated.

هاي عبارة مكملة لفوقها يعني مدام عندي ٧٥٪بدهم اقل من ٣٠ دقيقة ف انا عندي ٢٥٪ بدهم اكثر من ٣٠دقيقة

- 25% only of these tablets need more than 30 minutes to disintegrate.

Disintegration time (min)	Frequency
5	1
10	4
12	1
15	1
20	1
25	2
30	2
40	2
60	1
Total	15

Measures of Position

زوجي

- **Example (n is even)**

The time in minutes needed a random sample of 20 capsule to disintegrate are as follow:

يتفكك **(Q1)** **Median (Q2)** **(Q3)**

5 10 10 15 15 15 15 20 20 20 | 25 30 30 40 40 45 60 60 65 85

Find the quartiles Q1, Q2, and Q3?

هون زوجي بختلف عن فردي لانو ما حسبنا من ربع Q2=medain الاول و لا من ربع ثاني

1. Data is already in an order from smallest to highest
2. The **MD** is $X_{((n+1)/2)} = X_{(10.5)} = (20+25)/2 =$ **22.5** minutes.
3. For the half to the left of the median, n=10, then the first quartiles is **Q1**=(15+15)/2=**15** minutes
4. The second quartiles is Q2=MD=22.5 minutes.
5. For the half to the right with the MD, n=10, then the third quartile is **Q3**=(40+45)/2=**42.5** minutes.
6. The quartiles are: **Q1=15 minutes**, **Q2=22.5 minutes**, **Q3=42.5 minutes**.

Measures of Position

Conclusion

- This means that 25% of the capsules need less than 15 minutes to disintegrate.
- 50% of the capsules need 22.5 minutes to disintegrate.
- Before 42.5 minutes 75% of all capsules were disintegrated.
- 25% only of these capsules need more than 42.5 minutes to disintegrate.

Disintegration Time (minutes)	Frequency
5	1
10	2
15	4
20	3
25	1
30	2
40	2
45	1
60	2
65	1
85	1
Total	20

Measures of Position

- **Example**

The Gulf Pharmaceutical Industries Company in UAE planning to improve safety plan in its factory. For this, accident data for the last 50 weeks was compiled. These data were listed into simple frequency table as shown below:

Number of Accidents (x)	2	7	12	17	22
Number of weeks (f)	5	19	13	8	5

Find the quartiles Q1, Q2 and Q3 for number of accidents per week?

Measures of Position

- **Conclusion:**

The results mean that:

نفس شرح يلي فوق

- ❑ 25% of weeks have less than 7 accidents and 75% of weeks have more than 7 accidents.

- ❑ 50% of weeks have less than 12 accidents and 50% of weeks have more than 12 accidents.

- ❑ 75% of weeks have less than 17 accidents and 25% of weeks have more than 17 accidents.

Measures of Position

C. The quartiles for grouped frequency table

For the second quartile (Median) (Q2=MD)

Step-1: Construct the cumulative frequency distribution.

Step-2: Determine the second quartile class interval (**Second Quartile Class**),

that is, the first class having cumulative frequency (cf) greater than or equal to (n/2).

Step-3: Find the second quartile (Q2) by using the following formula:

$$Q_2 = L_2 + \left(\frac{\left(\frac{n}{2}\right) - cf_2}{f_2} \right) * i$$

Where:

n= the total number of frequencies.

cf_2= cumulative frequency prior to the second quartile class interval.

i = the class interval width.

L_2 = the lower boundary (limit) of the second quartile class interval.

f_2= the frequency of the second quartile class interval.

Measures of Position

C. The quartiles for grouped frequency table

For the third quartile (Q3)

Step-1: Construct the cumulative frequency distribution.

Step-2: Determine the third quartile class interval (**Third Quartile Class**),

that is, the first class having cumulative frequency (cf) greater than or equal to (3n/4).

Step-3: Find the third quartile (Q2) by using the following formula:

$$Q_3 = L_3 + \left(\frac{\left(\frac{3n}{4}\right) - cf_3}{f_3} \right) * i$$

Where:

n= the total number of frequencies.

cf_3= cumulative frequency prior to the third quartile class interval.

i = the class interval width.

L_3 = the lower boundary (limit) of the third quartile class interval.

f_3= the frequency of the third quartile class interval.

Measures of Position

- **Example:**

The following frequency table represents the daily sales volume in JD for a period of 30 days selected from the sales records of given pharmacy in Jordan:

Class (sales volume in JD)	53 - 56	57 - 60	61 - 64	65 - 68	69 - 72
Number of Days (f)	3	5	9	7	6

Find the quartiles Q_1, Q_2, and Q_3 for the sales volume in JD?

Measures of Position

- **Step-3: calculate the value of the Q1 a follows:**

n/4= 30/4 =7.5, then (57-60) is the first quartile class, then

$$Q_1 = L_1 + \left(\frac{\left(\frac{n}{4}\right) - cf_1}{f_1}\right) * i = 56.5 + \left(\frac{7.5-3}{5}\right) * 4 = 60.1 \text{ JD.}$$

- **Step-4: calculate the value of the Q2 a follows:**

n/2= 30/2 = 15, then (61-64) is the second quartile class (median class), then

$$Q_2 = L_2 + \left(\frac{\left(\frac{n}{2}\right) - cf_2}{f_2}\right) * i = 60.5 + \left(\frac{15-8}{9}\right) * 4 = 63.61 \text{ JD.}$$

- **Step-5: calculate the value of the Q3 a follows:**

3n/4=(3*30)/4= 22.5, then (65-68) is the third quartile class, then

$$Q_3 = L_3 + \left(\frac{\left(\frac{3n}{4}\right) - cf_3}{f_3}\right) * i = 64.5 + \left(\frac{22.5-17}{7}\right) * 4 = 67.64 \text{ JD.}$$

Measures of Position

Conclusion

For this pharmacy, the results mean that:

- 25% of days have sales volume less than 61.1 JD and 75% of days have sales volume more than 61.1 JD.

- 50% of days have sales volume less than 63.61 and 50% of days have sales volume more than 63.61 JD.

- 75% of days have sales volume less than 67.64 JD and 25% of days have sales volume more than 67.64 JD.

Measures of Position

Quartiles

$Q_1 = P_{25}$

$Q_2 = P_{50}$

$Q_3 = P_{75}$

P= Percentile

Interquartile Range (or IQR): $Q_3 - Q_1$

Semi-interquartile Range: $\frac{Q_3 - Q_1}{2}$

Midquartile Range: $\frac{Q_1 + Q_3}{2}$

10 - 90 Percentile Range: $P_{90} - P_{10}$

Measures of Position

The Interquartile Range (IQR)

مقياس قوي

- The interquartile range (IQR) is a robust measure of variation that is based on the quartiles. يعتمد ع ارباع

- The IQR is defined as the range of middle 50% of observations in the data set.

- It is the difference between the third quartile (Q3) and the first quartile (Q1), and it is found by using following formula:

$$IQR = Q3 - Q1$$

5 =2/(9+1)

Median = 72

Lower half | Upper half

63 64 64 70 72 76 77 81 81

Lower quarter | Upper quarter

Interquartile range: 79–64 = 15

$Q_1 = (64+64)/2=64$

$Q_3 = (77+81)/2=79$

Median = 71

Lower half | Upper half

62 63 64 64 70 72 76 77 81 81

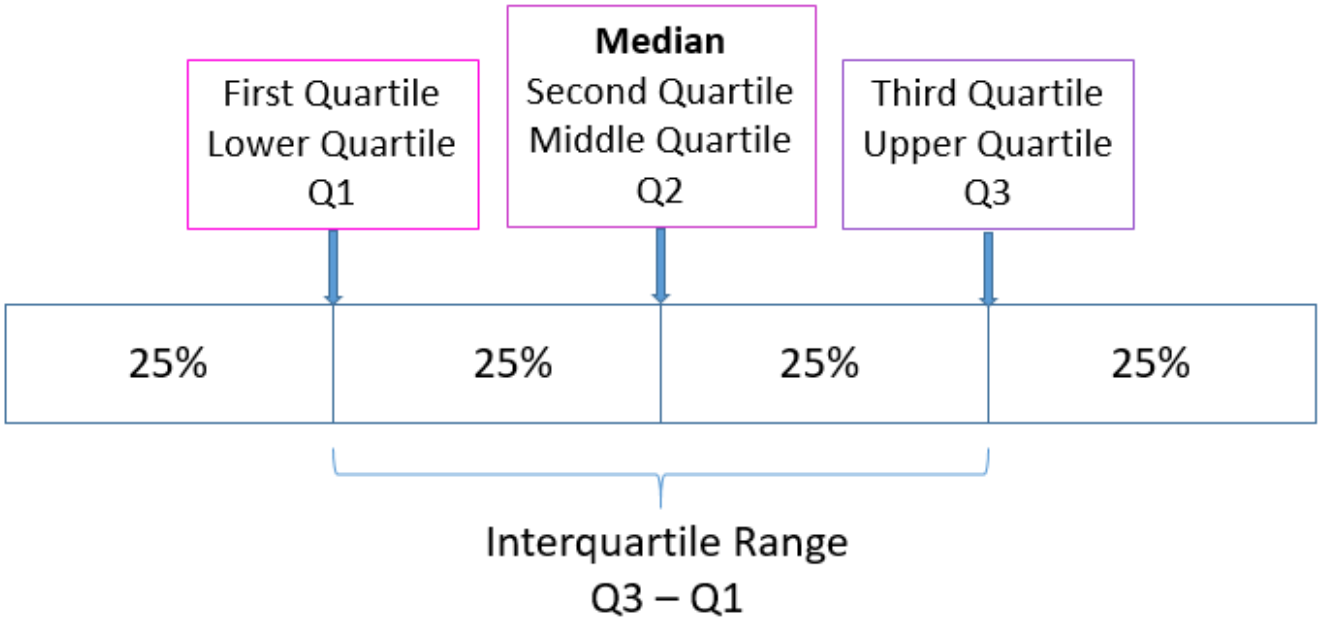
Lower quarter | Upper quarter

Interquartile range: 77–64 = 13

$Q_1 = 64$

$Q_3 = 77$

Median and Quartiles



Median and inter-quartile range are robust indicators of central tendency and dispersion

The median (second quartile) and inter-quartile range can be used as an alternative method for describing the central tendency and dispersion of a set of measured data. Both are robust and can be useful where there are occasional extreme values.

Measures of Position

- Example: we want to know the **IQR**, **Semi-interquartile Range**, **and the Midquartile** of the ordered grades from smallest to largest, **n=24**

Q_1 Median Q_2 Q_3

58,59,68,69,74,75,78,78,78,79,79,81,83,84,85,85,86,86,88,88,89,90,90,92

اول إشي بدي ارتب اذا ما كان مرتب

median $= Q_2$

$\frac{24+1}{2} = 12.5 = \frac{81+83}{2} = 82$

$Q_1 = \frac{12+1}{2} = 6.5 = \frac{75+78}{2} = 76.5$

$Q_3 = \frac{12+1}{2} = 6.5 = \frac{86+88}{2} = 87$

Measures of Position

Example: we want to know the Q1, Q2, Q3, IQR, Semi-interquartile Range , and the Midquartile, and the 10-90 percentile range of the ordered grades from smallest to largest, n=24

58,59,68,69,74,75,78,78,78,79,79,81,83,84,85,85,86,86,88,88,89,90,90,92

Q1=25%=25/100*24= 6 + 0.5 = 6.5 (position) which corresponds to 76.5.

Q2=50 / 100 *24= 12 +0.5 (position) which corresponds to 82.

Q3=75 / 100 *24= 18+0.5 (position) which corresponds to 87.

بنقدر نحلهم في علم تقييم

IQR= Q3–Q1= 87-76.5=10.5

Semi-interquartile Range= 10.5/2=5.25

Midquartile= Q1+Q3/2 = 81.75

10-90 percentile range=P90-P10.

P90= 0.9*24= 21.6+0.5=22=90, P10=0.1*24=2.4+0.5=2.9=68

10-90 percentile range=P90-P10= 90-68= 22