آمار مدرس: فاطمه عسگری

فصل یک :

آمار توصیفی: تعریف آمار » مجموعه ایی از تکنیک ها و روشهایی گفته می شود که در جمع آوری ، طبقه بندی ، خلاصه کردن و تجزیه و تحلیل و تفسیر اطلاعات مورد استفاده قرار می گیرد

آمار توصیفی آن بخش از آمار است که به جمع آوری ، خلاصه کردن ، نمایش و پردازش اطلاعات می پردازد بدون آنکه هرگونه نتیجه گیری از این اطلاعات انجام دهد.

آمار استنباطی : در این بخش از آمار اطلاعات بدست امده از آمار توصیفی مورد تجزیه و تحلیل قرارمیگیرد و براساس تحلیل های انجام شده نتیجه گیری و استنباط بعمل می آید که می توانیم این نتیجه گیری را به کل جامعه نیز تعمیم دهیم.

جامعه آماری :

عبارت است از مجموعه ای از افراد یا اشیائ یا عناصری که حداقل دریک صفت مشترک باشند تعداد عناصر موجود را حجم جامعه گوییم

اگر تعداد عناصر جامعه قابل شمارش و متناهی باشد جامعه را محدود در غیر اینصورت جامعه را نامحدود گوییم.

صفت:

کیفیت یا کمیتی است که متعلق به عناصر جامعه ای آماری هست .

صفت ثابت :

صفتی است که بین همه عناصر جامعه آماری مشترک است مثلا در جامعه ئ دانشگاه، دانشجو بودن.

صفت متغیر :

صفتی است که از افرادی به فرد دیگردر جامعه آماری تغییر می کند مثلا : سن و قددانشجویان

صفت متغیر کمی :

هرگاه اندازه صفتی را بتوان با ابزار های رایج اندازه گیری کرد و با یک عدد واحد دار بیان نمود ان را صفت متغیر کمی می نامیم.
(مثل : سن و قد و وزن) صفت متغیر کمی راگسسته هم گویند هرگاه اندازه عددی آن از طریق شمارش بدست آید مانند تعداد کارکنان یک سازمان صفت متغیر کمی را پیوسته گویند هرگاه از طریق اندازه گیری بدست آید مانند: قد و وزن ،مدت زمان

صفت متغیرکیفی :

هرگاه اندازه صفتی را نتوان با ابزارهای رایج اندازه گیری کرد و به صورت یه عدد واحد دار ان را صفت متغیر کیفی گوییم مانند مهارت
زیبایی،استعداد .صفت متغیر کیفی در جامعه یا ترتیبی یا اسمی است متغیر های کیفی که در آنها نوعی ترتیب طبیعی وجود دارد متغیر های
کیفی ترتیبی می نامند مانند مراحل زندگی.متغیر کیفی که ترتیبی نباشد را متغیر اسمی گوییم مانند رنگ چشم و گروه خونی .

نمونه :

بخش کوچکی از جامعه آماری است که به عنوان الگوی جامعه مورد تفسیر از نظر صفت مورد مطالعه با روشهای علمی انتخاب شده و باید گویای کاملی از وضعیت جامعه از نظر صفت مورد مطالعه باشد تعداد عناصر موجود در نمونه را حجم نمونه گوییم و با حرف N نشان می دهیم.

داده های آماری عبارت از اندازه های صفت متغیر عناصر جامعه یا نمونه اماری که با استفاده از اندازه گیری،آزمایش ،مشاهده و...بدست می آید.

تنظیم طبقه بندی داده ها :

سازماندهی مشاهدات را در امار توزیع فراوانی گوییم .توزیع فراوانی وسیله ای مناسب برای خلاصه کردن و مشخص نمودن ویژگی های اصلی داده های خام تحقیق است.

نوع هریک از متغیر های زیر را مشخص کنید.

1- دمای اتاق » متغیر کمی پیوسته

2 – غذاهای موجودر یک رستوران » کیفی اسمی

3- فشارهوا در بالای کوه » کمی پیوسته

4- گرایش سیاسی افراد » کیفی اسمی

5- مراحل بار دهی درخت میوه » کیفی ترتیبی

6- تعداد سئوالات امتحانی » کمی گسسته

جدول توزیع فراوانی

نحوه دسته بندی داده ها و اطلاعات و خلاصه کردن آنها جدول توزیع فراوانی است بر حسب نوع داده ها (گسسته ، پیوسته) نوع جدولها با هم فرق دارند.

جدول توزیع فراوانی برای داده های گسسته :

برای تنظیم داده های کیفی و داده های گسسته که تنوع اندک دارند از توزیع فراوانی طبقه بندی نشده استفاده می کنیم در این جدول در یک ردیف حالات مختلف متغیرو در ردیف دیگر تعداد اعضائ(فراوانی) هر حالت را می نویسیم

تعداد حالات گروه خونی	A	B	AB	O
	70	60	30	40

برای داده های گسسته و کیفی انجام می شود اگر تنوع زیاد نباشد.اگر تنوع تعداد حالات زیاد باشد این ها را هم باید طبقه بندی کنیم.

جدول توزیع فراوانی طبقه بندی شده:

برای تنظیم داده های پیوسته و حتی داده های گسسته ای که تنوع آنها زیاد باشد از این جدول استفاده می شود در این حالت اعداد را در طبقات جدا از هم دسته بندی کرده و مراحل زیر را انجام می دهیم.

مثال:مدیر یک موسسه تاکسی سرویس مایل است بداند اتومبیل های موسسه روزانه چند کیلومتر مسافت طی می کنند . برای این منظور کارکرد 24 اتومبیل را یادداشت کرده است .داده های زیر بدست امده.

138،107،136،128،148،118،142،129،115،123،133،121،128،122،144،123،126،135،125،98،117،153، 141،126

ابتدا کمترین داده (x_{min}) و بیشترین داده را (x_{max}) را در نظر می گیریم که با Xi نشان می دهیم .

N = تعداد داده ها x_{min}=98 x_{ma} =153

دامنه تغییرات را با R نشان می دهیم و تعداد دسته را با K نشان می دهیم.

دامنه تغییرات   R=55   R= $x_{max} - x_{min}$=153-98=55

k = 5   K = $[\sqrt{n}]$ = $[\sqrt{24}]$ = 5

$\frac{55}{5} = 11 = \frac{R}{K}$ طول دسته

برای رسم جدول کمترین داده را با طول دسته جمع کرده 109 = 11 + 98 حال جدول را رسم می نمائیم.

f_i تعداد اعداد بین دسته در داده های اولیه که همان n هست می باشد .مثلا بین اعداد 109 تا 98 تعداد 2 عدد در داده های اولیه موجود است

$L_i - L_{i+1}$				f_i %	
98 – 109	2	103/5			2
109 – 120	3	114/5			5
120 – 131	10	125/5			15
131 – 142	5	136/5			20
142 - 153	4	147/5			24

مرکز دسته $c_i = \frac{L_i + L_{i+1}}{2} = \frac{98+10}{2} = \frac{207}{2} = 103/5$

(جمع اولین عدد و اخرین عدد هر دسته تقسیم بر 2= c_i

$c_i = \frac{\text{حدپائین+حدبالا}}{2}$

فراوانی دسته های ماقبل + فراوانی هر دسته= تجمعی فراوانی f_{ci}

مثال : نمرات زبان 20 دانشجو در یک موسسه زبان به قرار زیر است جدول توزیع فراوانی داده های طبقه بندی شده را برا ی آنها رسم کنید.

66-60-89-92-72-68-73-82-83-68-59-57-64-90-85-68-52-75-90-70

$x_{\min} = 52$ $x_{max} = 92$

$R = x_{max} - x_{\min} = 52\text{-}92 = 40$ R=40

$K = [\sqrt{n}\,] = [\sqrt{20}\,] = 5$ K = 5

طول دسته$\frac{40}{5} = 8 = = \frac{R}{K}$

$L_i - L_{i+1}$				f_i %	
52–60	3	56			3
60–68	3	64			6
68–76	7	72			13
76–84	2	80			15
84 - 92	5	88			20

$c_i = \frac{52+60}{2} = 56$, $\overline{f_i} = \frac{f_i}{n} =$, $\%f_i = \overline{f_i} \times 100$

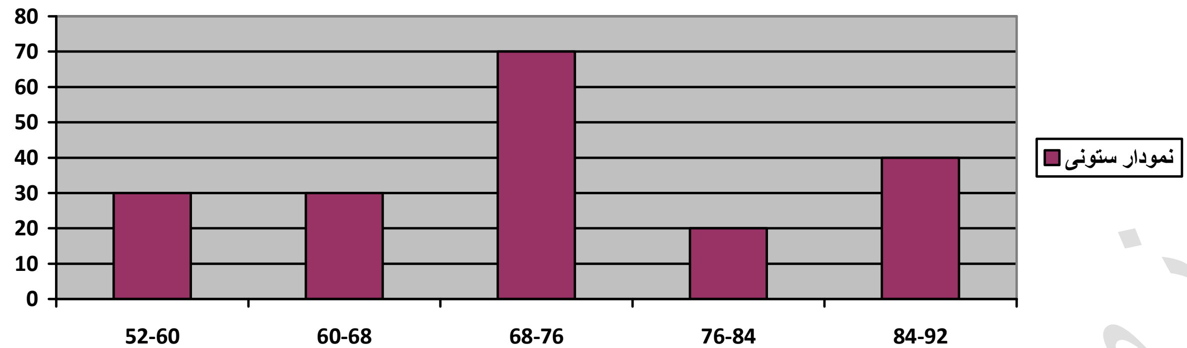
فراوانی دسته های ماقبل + فراوانی هر دسته= تجمعی فراوانی f_{ci}

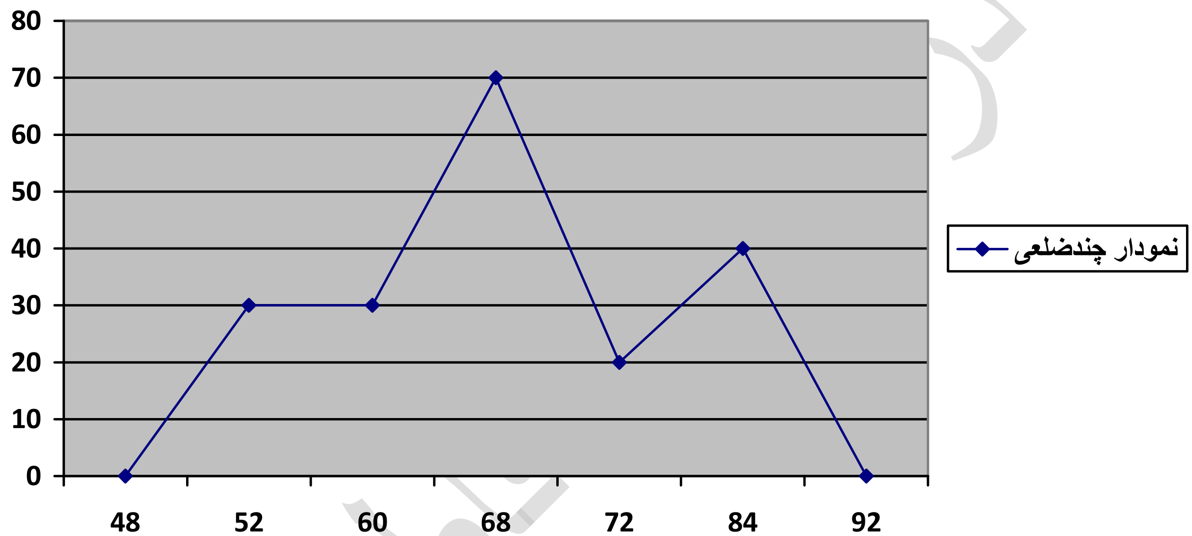
رسم نمودار جدول توزیع فراوانی داده های طبقه بندی شده:

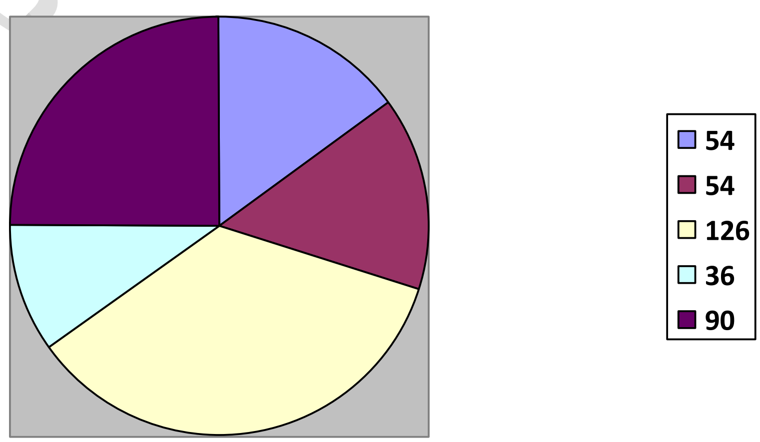
1-نمودار مستطیلی(ستونی) :ابتدا دو محور عمود بر هم رسم می کنیم. محور افقی محور داده ها و محور عمودی محور فراوانی می باشد برای هر دسته یک مستطیل که عرض آن همان دسته و ارتفاع آن فراوانی آن دسته باشد رسم می کنیم

2- نمودار چند ضلعی :ابتدا دو محور عمود بر هم رسم می کنیم. محور افقی محور داده ها و محور عمودی محور فراوانی می باشد برای هر دسته یک نقطه که مختصات آن مرکز دسته و فراوانی آن باشد رسم می کنیم سپس نقاط را به هم وصل می کنیم

3- نمودار دایره ای: : برای هر دسته یک زاویه از فرمول زیر به دست آورده و آن را روی دایره نشان می دهیم







$\alpha_1 = \frac{3}{20} \times 360 = 54$ $\alpha_2 = \frac{3}{20} \times 360 = 54$ $,\alpha_3 = \frac{7}{20} \times 360 = 126$ $,\alpha_4 = \frac{2}{20} \times 360 = 36,$

$\alpha_5 = \frac{5}{20} \times 360 = 90$

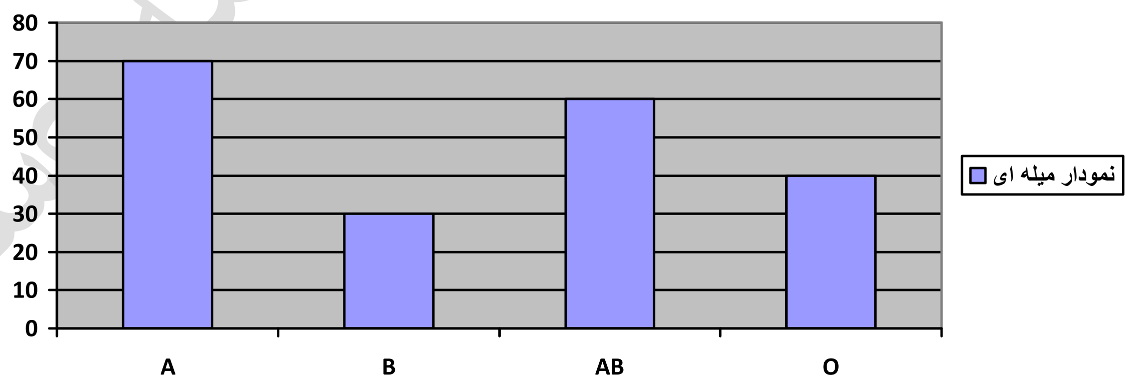
نمودارداده های گسسته یا کیفی

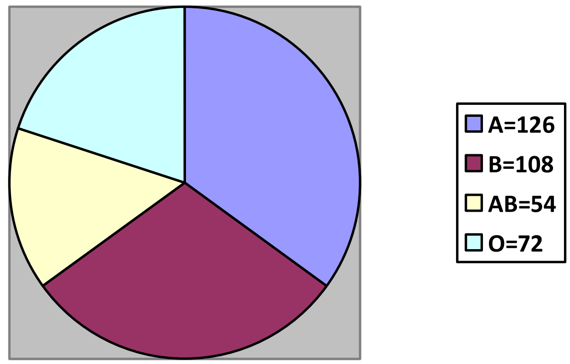
1-نمودار میله ای: ابتدا دو محور عمود بر هم رسم می کنیم. محور افقی محور داده ها و محور عمودی محور فراوانی می باشد برای هر داده یک میله که ابتدای آن همان داده و ارتفاع آن فراوانی آن داده باشد رسم می کنیم

2-نمودار دایره ای : برای هر داده یک زاویه از فرمول زیر به دست آورده و آن را روی دایره نشان می دهیم

مثال: نمودار های جدول زیر را رسم کنید؟

A	70
B	30
AB	60
O	40





شاخص :

برای اینکه داده ها را به طور عینی تر توصیف کنیم و مجموعه داده ها را با هم مقایسه کنیم از پارامتری استفاده می کنیم که تمام داده ها را در قالب یک عدد خلاصه کند به این عدد شاخص عددی گوییم.

انواع شاخص های عددی :

الف) شاخص های مرکزی که میزان گرایش به مرکز داده ها را بیان می کند.

ب) شاخص پراکندگی که میزان پراکندگی داده ها را بیان می کند.

شاخص های مرکزی بر سه نوع است:

1- **میانگین**   2- **میانه**   3- **مد**

میانگین :

به 3 دسته تقسیم می شود .   1- **حسابی**   2- **هندسی**   3- **هارمونیک**

میانگین حسابی داده های Xn . . . . . . X1 برابر است با $\bar{X} = \frac{x_1+x_2+x_3+.......x_n}{n}$

میانگین اعداد 20و12و5و3و8 را بدست آورید.

مثال : اگر 5و8و6و2 به ترتیب با فراوانی های 3و2و4و1 باشند میانگین را حساب کنید.

میانگین برای داده های پیوسته :

$L_i - L_{i+1}$		
2 – 6	4	4
6 – 10	3	8
10 – 14	2	12
14 - 18	4	16

میانگین هندسی (G) :

اگرداده های $x_1, x_2, x_3, \ldots\ldots, x_n$ را داشته باشیم

مثال : میانگین هندسی داده های 5 و 2 و 10 را حساب کنید.

$$G = \sqrt[3]{5 \times 2 \times 10} = 4/64$$

میانگین هارمونیک (H) :

فرمول میانگین هارمونیک :

$$H = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}$$

نکته : اگر مقیاس سنجش داده ها ترکیبی باشد مانند کیلومتر بر ساعت ، متر بر ثانیه و . . . برای یافتن متوسط انها از میانگین هارمونیک استفاده می کنیم .

مثال : فرض کنید کارگری یک کار معین را در 3 روز ، کارگر دوم در 4 روز و کار گر سوم در 6 روز تمام می کنند میانگین روزهای لازم برای اتمام اینکار در این کارگاه چند روز است؟

$$H = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}$$

مثال : اگر 3 اتومبیل کسیر 60 کیلومتری بین 2 منطقه را به ترتیب با سرعت 120 و 60 و 90 کیلومتر در ساعت طی نمایند میانگین سرعت این 3 اتومبیل چند کیلومتر بر ساعت است.؟

دومین شاخص مرکزی (مد یا نما) می باشد. m_o

داده ای که بیشترین فراوانی در بین داده ها را دارد مد گوئیم.

مثال : 12و 11و 12و 13 و14 و12 و 15 و13 عدد 12 مد این داده هاست

نکته : ممکن است در یک دسته 2 مد یا چند مد داشته باشیم و یا ممکن است اصلا مد نداشته باشیم مانند داده های 12 و13 و14 و15 و16 و17 که دارای مد نیستند.

مثال » 12و 11و 12و13 و14 و12 و 15 و13 و 13 $m_o = 12 , 13$

میانه :

عدد وسط داده ها را میانه گوییم برای بدست آوردن میانه ابتدا داده ها را از کوچک به بزرگ مرتب کرده اگر تعداد دادها فرد بود در این صورت عدد وسط داده ها میانه می باشد. در غیر این صورت اگر تعداد داده ها زوج بود 2 عدد وسط را بدست آورده میانگین این دو عدد برابر با میانه است.

مثال میانه : میانه داده های روبرو را پیدا کنید . 12و 12 و11 و13 و14 و12 و15 و13

ابتدا از کوچک به بزرگ مرتب می کنیم . 11و 12 و 12 و 12 و13 و 13 و14 و15 میانه 12.5 است

مثال : 11 و10 و8 و7 و5 و4 و2 » 2 و4 و5 و7 و8 و10 و11 » میانه این دسته عدد 7 می باشد.

میانگین داده های طبقه بندی شده:

مثال : میانگین جدول زیر را به دست آورید؟

$L_i - L_{i+1}$			
98 – 109	2	103/5	207
109 – 120	3	114/5	343/5
120 – 131	10	125/5	1/255
131 – 142	5	136/5	682/5
142 - 153	4	147/5	590

$$\frac{\sum f_i\, c_i}{\sum f_i} = \frac{590}{24} =$$

$$\bar{X} = \frac{3078}{24} = 128/25$$

$L_i - L_{i+1}$			

52 – 60	3	56	168
60 – 68	3	64	192
68 – 76	7	72	504
76 – 84	2	80	160
84 - 92	5	88	440

$$\bar{X} = \frac{\sum f_i\, c_i}{\sum f_i} = \frac{1464}{20} = \ 73/2$$

روش بدست آوردن میانه برای داده های طبقه بندی شده:

مثال :میانه داده های زیر را بدست اورید؟

$L_i - L_{i+1}$		
1 - 4	5	5
4 – 7	6	11
7 – 10	2	13
10 - 13	7	20

f_i = 6 , , n=20

f_{ci-1} =5 . L_i = 4 , L=3

ابتدا f_{ci} را بدست آورده و سپس وسط عدد n را بدست می اوریم $\frac{20}{2} = 10$ ، چون تعداد داده ها زوج است، عدد 10و11 عدد های وسط هستند، که دردسته دوم تجمعی فراوانی می باشند پس دسته ردیف دوم دسته میانه هست حال با جایگزینی اعداد در فرمول میانه را بدست می اوریم.

$$md = 4 + \frac{\frac{20}{2} - 5}{6} \times 3 = 4 + \frac{5}{6} \times 3 = 4 + \frac{5\times3}{6} = 4 + \frac{15}{6} = 4 + \frac{5}{2} = 4 + 2.5 = 6.5$$

مثال :میانه داده های زیر را بدست اورید.

f_i = 2, f_{ci-1} =11   , L=4   , n=27

$L_i - L_{i+1}$		
3 – 7	8	8
7 – 11	3	11
11 – 15	2	13
15 – 19	8	21
19 - 23	6	27

مثال:میانه داده های زیر را به دست آورید؟

$L_i - L_{i+1}$		
3 – 8	2	2
8 – 13	3	5
13 – 18	2	<u>7</u>
18 – 23	4	11
23 - 28	1	12

f_{ci-1} =5 f_i = 2   , n=12

. L_i = 13   , L=5

روش بدست آوردن مد برای داده های طبقه بندی شده :

مد :داده ی دارای بیشترین تکرار بین داده ها یعنی داده دارای بیشترین فراوانی f_i می باشد.

مثال :مد داده های زیر را به دست آورید.

$L_i - L_{i+1}$	
1 – 4	5
4 – 7	6
7 – 10	2
10 – 13	7
13 - 16	3

مثال : مد داده های زیر را به دست آورید؟

$L_i - L_{i+1}$	
3 – 7	7
7 – 11	3
11 – 15	2
15 – 19	8
19 - 23	5

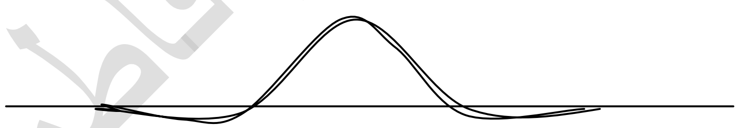
مثال : مد داده های زیر را به دست آورید.؟

$L_i - L_{i+1}$	
3 – 7	8
7 – 11	3
11 – 15	2
15 – 19	8

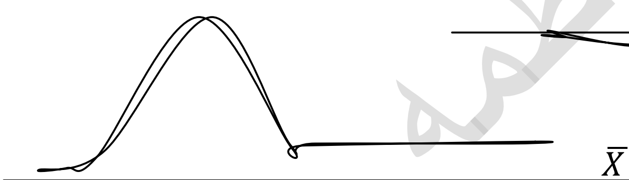
در این صورت که دو مد داریم برای هرکدام مجزا محاسبه می کنیم.

چولگی : در یک جامعه ای آماری ممکن است داده ها متقارن نباشند در اینصورت هرگونه عدم تقارن داده ها را چولگی گویند

الف) اگر داده ها متقارن باشند آنگاه$\bar{X} = m_o = md$میانه ، مد و میانگین با هم برابر باشند متقارن هستند

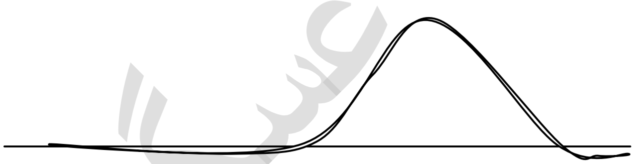


ب) اگر داده ها نامتقارن باشند 2 حالت داریم .



1- توزیع فراوانی داده ها چوله به سمت راست است $\bar{X} < m_o < md$

توزیع فراوانی داده ها چوله بسمت چپ است $\bar{X} > m_o > md$



مثال:مدیر یک موسسه تاکسی سرویس مایل است بداند اتومبیل های موسسه روزانه چند کیلومتر مسافت طی می کنند . برای این منظور کارکرد 24 اتومبیل را یادداشت کرده است .میانه و مد داده های زیرا بدست امده.

138،107،136،128،148،118،142،129،115،123،133،121،128،122،144،123،126،135،125،98،117،153،
126،141

ابتدا کمترین داده (x_{min}) و بیشترین داده را (x_{max}) را در نظر می گیریم.

$L_i - L_{i+1}$		
98 – 109	2	2
109 – 120	3	5
120 – 131	10	17
131 – 142	5	22
142 - 153	4	26

مثال : نمرات زبان 20 دانشجو در یک موسسه زبان به قرار زیر است جدول توزیع فراوانی داده های طبقه بندی شده را برای آنها رسم کنید.

66-60-89-92-72-68-73-82-83-68-59-57-64-90-85-68-52-75-90-70

$L_i - L_{i+1}$		
52 – 60	3	3
60 – 68	3	6
68 – 76	7	13
76 – 84	2	15
84 - 92	5	20

شاخص های پراکندگی: میزان پراکندگی و پراکنش داده ها را نشان می دهند

1- انحراف از میانگین

$$m_d = \frac{\sum_{i=1}^{n}|x_i - \overline{x}|}{n}$$

مثال: انحراف میانگین داده‌های زیر را بدست آورید؟

2- واریانس

مثال : واریانس داده‌های زیر را بدست آورید.

3- انحراف معیار (S)

$S = \sqrt{S^2}$

مثال 1) انحراف معیار (S) داده‌های زیر را بدست آورید .

3/28S = $\sqrt{10/8}$ =

تمرین :

شاخص های عددی برای داده های زیر محاسبه کنید . عدد شاخص را بدست آورید (پراکندگی – مرکزی) .

$$\bar{x} = \frac{4 \times 3 + 5 \times 2 + 1 \times 6 + 2 \times 1}{12} = \frac{12+10+6+2}{12} = \frac{30}{12} = 2/5$$

x i	f i
4	3
5	2
1	6
2	1
	n =
	12