بسم الله الرحمن الرحیم

توصیف اطلاعات:

شاخص های مرکزی

شاخص های مرکزی

مطالبی که در این جلسه خواهید آموخت

۱– معرفی شاخص ها

۲– میانگین ها

- میانگین حسابی
- میانگین کل
- میانگین هندسی
- میانگین همساز

۳– میانه

۴– نما (مد)

شاخص های مرکزی

گام اول در توصیف داده ها، طبقه بندی آنها است که در پیشتر به آن پرداخته شد. برای اینکه داده‌ها را به صورت بهتری توصیف کنیم، باید گام دیگری برداریم و آن اینکه مشخص کننده‌های عددی را بـرای داده هـا بدسـت آوریـم. بـه ایـن مشخص‌کننده‌های عددی شاخص یا معیار می‌گویند و بر دو نوعند:

شاخص های مرکزی: شاخص‌هایی هستند که میزان گرایش به مرکز داده ها را اندازه می‌گیرند.

شاخص های پراکندگی: شاخص‌هایی هستند که میزان پراکندگی داده ها از مرکز را اندازه می‌گیرند.

توجه: بسیاری از شاخص های آماری برای داده‌های کیفی تعریف نمی‌شوند

شاخص های مرکزی

- میانگین
- میانه
- نما
- چندک ها

شاخص های پراکندگی

- واریانس
- انحراف معیار
- دامنه تغییرات
- ضریب تغییرات
- ضریب چولگی

شاخص های مرکزی

میانگین حسابی

میانگین حسابی

یکی از مهمترین شاخص های گرایش به مرکز، میانگین حسابی است. میانگین حسابی را از رابطه زیر به دست می آوریم. چنانچه میانگین حسابی را از یک نمونه به دست آوریم آن را با $\overline{x}$ و اگر میانگین مقادیر مربوط به همه جمعیت را محاسبه کنیم،
آن را با μ (مو) نمایش می‌دهیم.

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

مثال: فرض کنید داده‌های زیر، نمرات ۱۰ درس مربوط به یک دانش‌آموز باشد. می‌خواهیم میانگین نمرات او را حساب کنیم، داریم:

۱۱ ۱۰ ۱۶ ۱۲ ۱٤ ۱۳ ۱٥ ۱٤ ۹ ۱٦

$$\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}(11+10+16+12+14+13+15+14+9+16) = \frac{130}{10} = 13$$

توجه داشته باشید :

اگر داده‌ها طبقه بندی شده باشند یعنی مقادیر $x_1, x_2, \ldots x_k$ را با فراوانی های $f_1, f_2, \ldots, f_k$ در اختیار داشته‌باشیم، در این صورت فرمول میانگین به شکل زیر خواهد بود:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i = \frac{1}{n}(f_1x_1 + f_2x_2 + \ldots + f_kx_k)$$

شاخص های مرکزی

میانگین حسابی

مثال: میانگین حسابی را در داده‌های جدول زیر محاسبه می کنیم

x_i	f_i	$f_i x_i$
5	3	3×5=15
7	3	3×7=21
8	2	2×8=16
9	5	5×9=45
10	5	5×10=50
	18	$\sum f_i x_i = ۱۴۷$

$$\bar{x} = \frac{۱}{۱۸}\sum_{i=۱}^{۵} f_i x_i = \frac{۱}{۱۸}(۱۴۷) = ۸/۱۷$$

توجه داشته باشید

در داده های پیوسته که طبقه بندی شده باشند، در فرمول بالا باید به جای x_i از x'_i (نماینده طبقات) استفاده کنیم.

مثال: برای داده های جدول توزیع فراوانی مقابل میانگین حسابی را به دست می آوریم.

حدود طبقات	f_i	x'_i	$f_i x'_i$
1-3	1	2	2
4-6	5	5	25
7-9	15	8	120
10-12	6	11	66
13-15	3	14	42
	30		255

$$\bar{x} = \frac{۱}{۳۰}\sum_{i=۱}^{۵} f_i x'_i = \frac{۱}{۳۰}(۲۵۵) = ۸/۵$$

شاخص‌های مرکزی

میانگین حسابی

خواص میانگین حسابی

۱- همیشه مجموع اختلاف داده‌ها از میانگین حسابی صفر است.

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

۲-هرگاه به هر یک از داده‌ها عدد ثابتی را مانند a اضافه یا کم کنیم، میانگین داده‌های حاصل، برابر با میانگین داده‌های قبلی به اضافه (یا منهای) عدد ثابت a خواهد بود.

$$Y_i = X_i \pm a \rightarrow \bar{Y} = \bar{X} \pm a$$

۳- هرگاه عدد ثابت a را در هریک از داده‌ها ضرب کنیم، میانگین داده‌های حاصل، برابر میانگین داده‌های قبلی ضرب در عدد ثابت a خواهد بود.

$$Y_i = aX_i \Rightarrow \bar{Y} = a\bar{X}$$

۴- درمورد تقسیم هم به صورت ضرب عمل می کنیم.

$$Y_i = \frac{X_i}{a} \rightarrow \bar{Y} = \frac{\bar{X}}{a}$$

شاخص های مرکزی

میانگین کل

اگر یک مجموعه n1 تائی از داده‌ها با میانگین $\bar{X}_۱$ و یک مجموعه n2 تائی از داده‌ها با میانگین $\bar{X}_۲$ داشته باشیم، میانگین کل این دو مجموعه داده برابر است با :

$$\bar{X} = \frac{n_i \bar{X}_۱ + n_۲ \bar{X}_۲}{n_۱ + n_۲}$$

مثال: در یک مجموعه، ۱۰ داده با میانگین $\bar{x}_۱ = ۴/۲$ و در مجموعه دیگری ۸ داده با میانگین $x_۲ = ۴$ در اختیار است. میانگین کل داده‌ها را به دست می آوریم.

$$\bar{x} = \frac{۴/۲ \times ۱۰ + ۴ \times ۸}{۱۰ + ۸} = \frac{۷۴}{۱۸} = ۴/۱۱$$

این موضوع برای n مجموعه نیز صادق است

کاربرد این میانگین بیشتر برای محاسبه حد متوسط شاخص‌ها، نسبت ها و درصدها است.

میانگین هندسی

فرض کنید x_1 , x_2 ,... x_n داده‌های مثبت و مخالف صفر باشند،

میانگین هندسی (Geometric Mean) آنها از فرمول زیر بدست می‌آید و آن را با G نمایش می‌دهیم:

$$G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \times x_2 \times \dots \times x_n}$$

میانگین هندسی اعداد زیر را حساب کنید:

1, 2, 4

$$G = \sqrt{1 * 2 * 4} = \sqrt[3]{8} = 2$$

میانگین هندسی

شاخص های مرکزی

مثال: جدول زیر سرمایه هر سال در یک فروشگاه را نمایش می‌دهد با توجه به مقادیر جدول، نرخ رشد سرمایه (میانگین نسبت رشد) را محاسبه می کنیم.

سال	سرمایه	نسبت رشد
1373	15	-
1374	20	$\frac{۲۰}{۱۵}=\frac{٤}{۳}$
1375	30	$\frac{۳۰}{۲۰}=\frac{۳}{۲}$
1376	30	$\frac{۳۰}{۳۰}=۱$
1377	35	$\frac{۳٥}{۳۰}=\frac{۷}{٦}$
1378	40	$\frac{٤۰}{۳٥}=\frac{۸}{۷}$

$$G=\sqrt[5]{\frac{4}{3}*\frac{3}{2}*1*\frac{7}{6}*\frac{8}{7}}=\sqrt[5]{2.66}$$

روش کوتاهتر این است که حجم تولید در سال آخر یعنی 40 را برحجم تولید در سال اول یعنی 15 تقسیم کنیم و سپس ریشه (n - 1) ام را به دست آوریم.

برای سهولت در محاسبه میانگین هندسی می‌توان از خواص لگاریتم استفاده کرد. بدین ترتیب که اگر از طرفین فرمول آن، لگاریتم بگیریم می توان از فرمول زیر که مشابه میانگین حسابی است، استفاده کرد.

$$\log G=\frac{۱}{n}\sum_{i=۱}^{n}\log x_i$$

شاخص های مرکزی

میانگین همساز

فرض کنید داده‌های $x_1, x_2, \ldots, x_n$ مخالف صفر باشند، میانگین همساز (هارمونیک) آنها از فرمول زیر بدست می‌آید و آن را با ***H*** نمایش می‌دهیم:

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

این میانگین بیشتر در مطالعه شبکه های برق، اپتیک و برخی نسبت ها استفاده می شود.

مثال: فرض کنید یک اتومبیل مسیر ٦٠ کیلومتری را با سرعت ٤٠ کیلومتر طی کرده و همین مسیر را با سرعت ٦٠ کیلومتر برگشته است.
متوسط سرعت در طی این ١٢٠ کیلومتر چقدر است؟

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{2}{\frac{1}{40} + \frac{1}{60}} = 48$$

توجه داشته باشید

اگر از میانگین حسابی متوسط سرعت را حساب کنیم $\bar{x} = \frac{1}{2}(40+60) = 50$ را به دست خواهیم آورد که عدد درستی نیست. زیرا زمان رفتن این اتومبیل یک ساعت و زمان برگشتن او یک ساعت ونیم طول کشیده است. پس او مسیر ١٢٠ کیلومتری را در مدت دو ساعت و نیم طی کرده که به طور متوسط سرعتی معادل ٤٨ کیلومتر داشته است.

شاخص های مرکزی

میانه

میانه؛ اندازه ای است که در وسط داده های آماری مرتب (صعودی/نزولی) قرار می گیرد.

به عبارت دیگر داده‌ای از میان داده‌ها که دقیقاً در وسط قرار دارد.

توجه داشته باشید

اگر تعداد داده‌ها فرد باشد، داده ای که در وسط قرار می گیرد، میانه است. ولی اگر تعداد داده‌ها زوج باشد، میانگین دو عدد وسط، میانه خواهد بود.

مثال: میانه مجموعه داده‌های زیر را بدست می آوریم.
۶ ۳ ٤ ۱ ۲ ۳ ۱ ۲ ۷ ۰ ۱ ۵ ۳ ۰ ۳

ابتدا داده‌ها را مرتب می کنیم.

۷ ٦ ۵ ٤ ۳ ۳ ۳ ۳ ۲ ۲ ۱ ۱ ۱ ۰ ۰

چون تعداد داده ها ۱۵ و عددی فرد است، داده هشتم میانه است زیرا:

$$۱۵\times\frac{۱}{۲}=۷/۵\cong۸ \rightarrow Md=۳$$

شاخص های مرکزی

میانه

مثال: برای داده‌های زیر میانه را محاسبه می‌کنیم.

١ ٣ ٧ ٥ ٩ ١ ٨ ٦ ٤ ٧ ١ ١ ٩ ٩ ٢ ٢

ابتدا داده‌ها را مرتب می‌کنیم.

١ ١ ١ ١ ٢ ٢ ٣ ٤ ٥ ٦ ٧ ٧ ٨ ٩ ٩ ٩

چون $n = ١٦$ عددی زوج است، پس معدل داده هشتم و نهم را به عنوان میانه در نظر می گیریم:

١ ١ ١ ١ ٢ ٢ ٣ ٤ ٥ ٦ ٧ ٧ ٨ ٩ ٩ ٩

$$Me = \frac{٤ + ٥}{٢} = ٤ / ٥$$

توجه کنید که:

در دو مثال اخیر داده ها گسسته و دسته بندی نشده اند

شاخص های مرکزی

میانه

اگر داده ها گسسته و دسته بندی هستند مراحل زیر را دنبال کنید

۱- ستون فراوانی تجمعی را بدست می آورید.

۲- مقدار $\frac{۱}{۲} \times n$ را بدست آورده و از ستون فراوانی تجمعی، طبقه میانه‌دار را معلوم کنید.

۳- مقدار متغیر آن طبقه، میانه داده ها است.

مثال: برای داده های جدول زیر میانه را به دست آورید.

xi	f	cf	
3	2	2	
5	5	7	
7	4	11	طبقه میانه دار
9	3	14	
12	3	17	
	17		

$$17 * \frac{1}{2} = 8.5$$

شاخص های مرکزی

میانه

اگر داده ها پیوسته و دسته بندی شده هستند مراحل زیر را دنبال کنید

۱- ستون فراوانی تجمعی را بدست می آورید.

۲- مقدار $\frac{۱}{۲} \times n$ را بدست آورده و از ستون فراوانی تجمعی، طبقه میانه‌دار را معلوم کنید.

۳- میانه را از فرمول زیر به دست آورید.

$$Md = L_i + \frac{\frac{n}{۲} - F_{i-۱}}{f_i} \times C$$

در این فرمول داریم:

$L_i =$ کران پائین طبقه‌ای است که میانه در آن قرار دارد.

$n =$ تعداد داده‌ها

$F_{i-۱} =$ فراوانی تجمعی طبقه قبل از طبقه‌ای که شامل میانه است.

$f_i =$ فراوانی مطلق طبقه‌ای است که میانه در آن قرار دارد.

$c =$ فاصله طبقات

میانه

مثال: میانه را برای داده های جدول زیر به دست آورید.

حدود طبقات	f_i	cf
1 – 4	8	8
4 – 7	5	13
7 – 10	13	26
10 – 13	15	41
13 - 16	9	50

فراوانی تجمعی

طبقه میانه دار

ابتدا ستون فراوانی تجمعی را تشکیل می‌دهیم
سپس داریم

$$۵۰\times\frac{۱}{۲}=۲۵$$

بنا بر این بیست و پنجمین عدد در طبقه سوم قرار دارد
حال باید از فرمول استفاده کنیم داریم:

$$Md=۷+\frac{۲۵-۱۳}{۱۳}\times۳=۷+۲/۷۷\Rightarrow Md=۹/۷۷$$

فاصله طبقات

$$Md=L_i+\frac{\frac{n}{۲}-F_{i-۱}}{f_i}\times C$$

شاخص‌های مرکزی

نمـــــا

نما یا مد (Mode) داده‌ای است که بیشتر از سایر داده‌ها تکرار شده باشد و آن را با نماد Mo نمایش می‌دهیم.

مثال: نما در داده‌های زیر عدد ۳ است. زیرا بیشتر از همه، عدد ۳ تکرار شده است،

3 1 5 9 2 3 1 7 3 2 3 4 3 3

اگر داده‌ها گسسته و طبقه بندی شده باشند مانند مثال زیر عمل کنید

مثال: نما در داده‌های جدول زیر، عدد ۱۵ است. چون فراوانی عدد ۱۵ بیشتر است.

x_i	f_i	
10	5	
15	19	
20	13	
25	8	

نمـــــا

نما در داده‌های طبقه بندی شده پیوسته

اگر داده ها پیوسته و طبقه‌بندی شده هستند، به صورتی که در زیر توضیح داده می‌شود، نما را محاسبه کنید.

ابتدا طبقه‌ای را که فراوانی آن ماکزیمم است معلوم کنید (طبقه نمادار)

سپس از فرمول مقابل برای محاسبه نما استفاده کنید .

$$Mo = l_i + \frac{D_1}{D_1 + D_2} \times C$$

که در این فرمول:

l_i = کران پائین طبقه‌ای است که فراونی مطلق آن ماکزیمم است.(طبقه نمادار)

D_1 = تفاضل فراوانی مطلق طبقه نمادار با طبقه قبل از آن است، یعنی $D_1 = f_i - f_{i-1}$

D_2 = تفاضل فراوانی مطلق طبقه نمادار با طبقه بعد از آن است. یعنی $D_2 = f_i - f_{i+1}$

C = فاصله طبقات

شاخص های مرکزی

نمـــــا

مثال: برای داده‌های جدول فراوانی زیر نما را محاسبه می کنیم.

چون طبقه چهارم بیشترین فراوانی را دارد، پس طبقه نمادار است. حال از فرمول استفاده می کنیم داریم:

$$Mo = l_i + \frac{D_1}{D_1 + D_2} \times C$$

حدود طبقات	f_i
1 – 4	8
4 – 7	5
7 – 10	13
10 – 13	15
13 – 16	9

$D_1 = ۱۵ - ۱۳ = ۲$

$D_2 = ۱۵ - ۹ = ۶$

فاصله طبقات

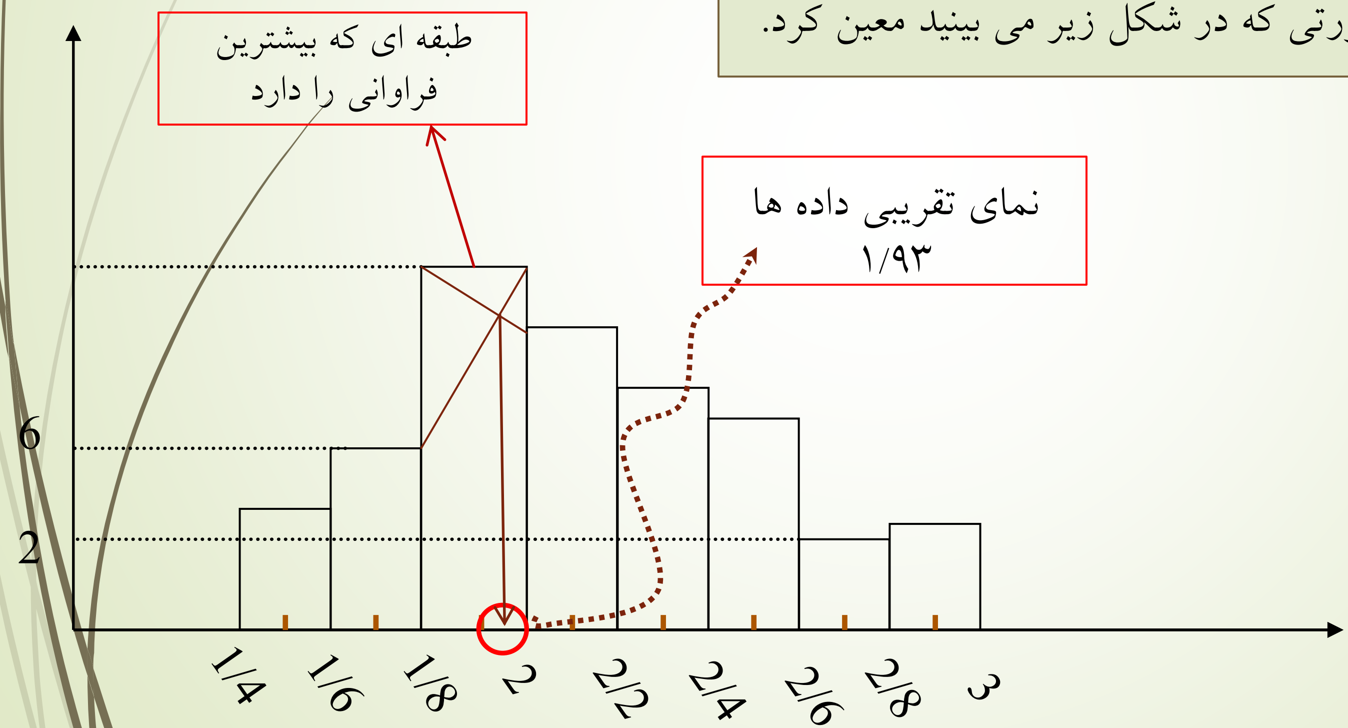
$$Mo = ۱۰ + \frac{۲}{۲ + ۶} \times ۳ = ۱۰ + ۰/۷۵ \Rightarrow Mo = ۱۰/۷۵$$

شاخص های مرکزی

نما

نما روی نمودار مستطیلی

می توان نمای داده ها را با استفاده از نمودار مستطیلی و به طور تقریبی و به صورتی که در شکل زیر می بینید معین کرد.



خودآزمایی:

باتوجه به مباحث درس اول و دوم، لطفا به پرسش های زیر پاسخ دهید

- ۱- کدام یک از مقیاس های زیر صفر مطلق دارد؟

  ۱) نسبی ۲) اسمی ۳) فاصله ای ۴) رتبه ای

- ۲- کدام مقیاس از ویژگی‌های بهتری برای اندازه‌گیری برخوردار است؟

  ۱) نسبی ۲) فاصله ای ۳) اسمی ۴) رتبه ای

- ۳- وزن محصولات یک کارخانه دارای چه نوع مقیاسی است؟
- ۱) رتبه ای ۲) نسبی ۳) اسمی ۴) فاصله ای

خودآزمایی:

باتوجه به مباحث درس اول و دوم، لطفا به پرسش های زیر پاسخ دهید

- ۴- جامعه آماری و نمونه را تعریف کنید.
- ۵- متغیر پیوسته و گسسته را با ذکر مثال، تعریف کنید.
- ۶- پارامتر چیست؟ چه تفاوتی با آماره دارد؟
- ۷- مقیاس های اندازه گیری اسمی، ترتیبی، فاصله ای و نسبی را با هم مقایسه نمایید.
- ۸- میانه و نما را در مجموعه داده زیر بدست آورید.

۶ ۳ ۴ ۱ ۲ ۳ ۱ ۲ ۷ ۰ ۱ ۵ ۳ ۰ ۳

۹- میانگین حسابی را در داده‌های جدول زیر محاسبه و نمودار هیستوگرام آن را ترسیم کنید.

x_i	f_i
۵	۳
۷	۳
۸	۲
۹	۵
۱۰	۵