

فصل چهار : داده های آماری

درس اول : گرد آوری داده ها

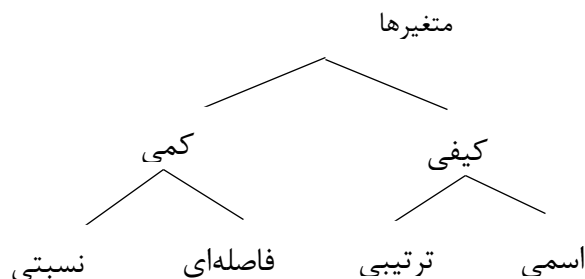
- داده‌ها واقعیت‌هایی درباره یک چیز هستند که در محاسبه، استنباط یا برنامه‌ریزی به کار می‌روند.
- واحد آماری به هر یک از افراد یا چیزهایی می‌گویند که داده‌های مربوط به آنها در یک بررسی آماری گردآوری می‌شود.
- مجموعه‌ی کل واحدهای آماری را جامعه‌ی آماری می‌گویند.
- هر زیر مجموعه از جامعه‌ی آماری را که با روش مشخصی انتخاب شده باشد، یک نمونه می‌نامند.
- نمونه‌ای که در آن همه اعضای جامعه شانس انتخاب یکسان در نمونه را داشته باشند، نمونه تصادفی می‌نامند.
- **آمارگیری** : گردآوری داده‌ها به یکی از روش‌های ممکن
- **آمارگر** : کسی که آمارگیری را انجام می‌دهد.

روش‌های گردآوری داده‌ها :

- الف) **مشاهده** : گردآوری داده‌ها بدون نیاز به فرد پاسخگو
- ب) **پرسشنامه** : مجموعه‌ی سوالات از پیش تعیین شده که توسط تعدادی پاسخ‌دهنده تکمیل می‌شود.
- پ) **مصاحبه** : معمولاً بین دو نفر صورت می‌گیرد که یکی مصاحبه‌گر و دیگری مصاحبه‌شونده است.
- ت) **دادگان‌ها** : شامل مجموعه‌ای از اطلاعات ذخیره شده هستند.
- تعریف متغیر** : هر ویژگی از اشخاص یا اشیاء که قرار است مورد بررسی قرار گیرد. متغیر نوع است.
- الف) **متغیر کمی** : متغیرهایی هستند که مقادیر عددی می‌گیرند و عملیات ریاضی برای آنها تعریف شده است.
- ب) **متغیر کیفی** : متغیرهایی هستند که برای دسته‌بندی افراد یا اشیاء در گروه‌ها به کار می‌روند.
- **پارامتر جامعه** : یک مشخصه عددی است که توصیف‌کننده‌ی جنبه‌ای خاص از جامعه است.
- **آمار نمونه** : مشخصه‌ای عددی که توصیف‌کننده‌ی جنبه خاص از نمونه است.

انواع متغیرها :

اسمی	این مقیاس برای متغیرهایی است که نام‌ها، برچسب‌ها و گروه‌ها هستند. در اینجا هیچ معیاری برای مرتب کردن آنها از کوچک به بزرگ وجود ندارد. مانند گروه خون
ترتیبی	این مقیاس برای متغیرهایی است که قابل مرتب کردن هستند در عین حال محاسبه اختلاف بین مقادیر داده شده یا امکان پذیر نیست یا بی‌معناست. مانند مدرک تحصیلی
فاصله‌ای	این مقیاس برای داده‌هایی است که قابل مرتب کردن هستند و همچنین اختلاف بین مقادیر داده‌ها با معنی است. مانند دمای هوا
نسبتی	این مقیاس برای داده‌هایی است که قابل مرتب کردن هستند و اختلاف بین مقادیر داده‌ها و نسبت مقادیر داده‌ها نیز با معناست. مانند قد و وزن



تعریف علم آمار: به مطالعه‌ی نحوه گردآوری، سازماندهی، تحلیل و تغییر داده‌ها جهت استخراج اطلاعات و تصمیم‌گیری، آمار گفته می‌شود.

درس دوم: معیارهای گرایش به مرکز

اگر x_1, x_2, \dots, x_n داده باشند، میانگین آن‌ها را با نماد \bar{x} نمایش می‌دهند و به صورت روبرو تعریف می‌شود.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

مثال: میانگین اعداد بین ۲۰ تا ۳۰ را بدست آورید.

$$\bar{x} = \frac{۲۱ + ۲۲ + ۲۳ + \dots + ۲۹}{۹} = ۲۵$$

داده‌ی دور افتاده: مقداری متفاوت با سایر مقادیر داده‌ها دارد. معمولاً مقدار آن بسیار بزرگ‌تر یا بسیار کوچک‌تر است.

میانه: داده‌ی وسطی داده‌های مرتب شده می‌باشد.

نکته: خواسته اصلی از آمار یک استنباط نیست بلکه در بررسی آماری دنبال برنامه‌ریزی و تصمیم‌گیری هستیم.

درس سوم: معیارهای پراکندگی

یک معیار معمول پراکندگی که نشان دهنده‌ی تغییرات داده‌های مورد مطالعه است انحراف استاندارد نام دارد.

اگر x_1, x_2, \dots, x_n داده باشند، انحراف استاندارد آن‌ها با نماد σ نشان داده و به صورت زیر تعریف می‌شود.

$$\sigma = \frac{1}{\sqrt{n}} \sqrt{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}$$

توان دوم انحراف استاندارد را واریانس می‌نامند.

$$\left. \begin{aligned} &(\bar{x} - 3\sigma, \bar{x} + 3\sigma) \\ &(\bar{x} - 2\sigma, \bar{x} + 2\sigma) \\ &(\bar{x} - \sigma, \bar{x} + \sigma) \end{aligned} \right\}$$

نکته: تقریباً ۹۹٪ مشاهدات بین ۳ برابر انحراف استاندارد از میانگین قرار دارند.

تقریباً مشاهدات بین ۲ برابر انحراف استاندارد از میانگین قرار دارند.

تقریباً مشاهدات بین یک برابر انحراف استاندارد از میانگین قرار دارند.

اگر داده‌ای دور افتاده وارد جامعه‌ی آماری شود، این کار سبب درهم ریختگی میانگین و انحراف استاندارد خواهد شد. در این زمان بهتر است از میانه (دامنه‌ی میان چارکی) استفاده شود. در اینجا معیار پراکندگی که تعریف می‌شود دامنه‌ی میان چارکی نامیده می‌شود (IQR) میانه وسط داده‌هاست و چارک اول یک چهارم مشاهدات و چارک سوم، سه چهارم مشاهدات خواهد بود.

علت استفاده از میانه و دامنه میان چارکی تغییر ساده و سریع آنهاست.

آزمون فصل چهارم

۲	<p>درستی یا نادرستی عبارت های زیر را مشخص کنید.</p> <p>الف) مجموعه کل واحد های آماری را جامعه آماری می گویند.</p> <p>ب) مصاحبه یک روش گرد آوری داده ها است.</p> <p>پ) پارامتر جامعه یک متغیر کمی است.</p> <p>ت) دور افتاده مقداری متفاوت با سایر مقادیر داده ها است.</p>	۱
۲/۵	<p>هر یک از جملات زیر را با اعداد یا کلمات مناسب کامل کنید.</p> <p>الف) نشان دهنده تغییرات داده های مورد مطالعه است.</p> <p>ب) تقریباً درصد مشاهدات بین یک انحراف معیار از میانگین است.</p> <p>پ) معیار پراکندگی که معمولاً با میانگین بیان می شود است.</p> <p>ت) متغیر ها به دو نوع و تقسیم می شوند.</p>	۲
۲	<p>گزینه ی مناسب را انتخاب کنید.</p> <p>(A) گروه خونی انسان چه نوع متغیری است؟</p> <p>الف) کیفی اسمی (ب) کیفی ترتیبی (پ) کمی فاصله ای (ت) کمی نسبتی</p> <p>(B) در داده های آماری ۲۲ و ۲۸ و ۲۵ و ۳۲ و ۳۰ و ۱۲ و ۴۰ میانه کدام است؟</p> <p>الف) ۲۵ (ب) ۲۸ (پ) ۲۲ (ت) ۳۰</p> <p>(C) توان دوم انحراف معیار نام دارد.</p> <p>الف) چارک (ب) میانگین (پ) میانه (ت) واریانس</p> <p>(D) نسبت مردان در کل جامعه کوهنوردان معرف یک است.</p> <p>الف) آماره (ب) پارامتر (پ) مقیاس (ت) متغیر</p>	۳
۱	<p>در هر مورد نوع داده ها را مشخص کنید.</p> <p>شماره داوطلبی رتبه های یک مسابقه وزن دانش آموزان دمای شهرها</p>	۴

۱/۵	در یک مطالعه از ۲۱۲۰ خانواده صرفه جو سوال شده است که کدام مورد صرفه جویی (آب - برق - گاز) بهتر است؟ الف) متغیر را مشخص کنید. ب) این متغیر کمی است یا کیفی؟ پ) جامعه آماری در اینجا چیست؟	۵																				
۲	در داده های آماری زیر میانگین و میانه را مشخص کنید. ۴۱ و ۳۹ و ۴۰ و ۲۹ و ۶۳ و ۵۰ و ۲۰ و ۱۹ و ۵۱	۶																				
۲	در داده های آماری زیر میانگین ، انحراف معیار و واریانس را بدست آورید. ۴۹ و ۹۱ و ۸۷ و ۶۷ و ۷۱ و ۴۵ و ۳۹ و ۵۷ و ۷۹ و ۸۵	۷																				
۲	یک موضوع برای هر یک از روش های جمع آوری داده ها بیان کنید. از طریق مصاحبه از طریق مشاهده از طریق دادگان از طریق پرسشنامه	۸																				
۳/۵	می خواهیم مدت زمانی که یک کودک به تماشای تلویزیون در طول مدت یک ماه صرف می کند را آمارگیری کنیم . الف) در این آمار گیری جامعه آماری را مشخص کنید. ب) یک روش نمونه گیری برای انتخاب نمونه معین کنید. پ) از چه روشی برای گرد آوری اطلاعات استفاده می کنید؟ ت) اندازه (تعداد) این جامعه چقدر است؟ ث) متغیر مورد مطالعه چیست؟ ج) متغیر مورد مطالعه از چه نوعی است؟	۹																				
۱	دو دسته داده زیر در کلاس الف و ب در درس ریاضی دهم می باشند. <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>الف</td> <td>۵</td> <td>۱۱</td> <td>۱۹</td> <td>۱۷</td> <td>۱۴</td> <td>۱۳</td> <td>۱۰</td> <td>۱۱</td> <td>۲۰</td> </tr> <tr> <td>ب</td> <td>۹</td> <td>۷</td> <td>۱۸</td> <td>۱۳</td> <td>۱۹</td> <td>۱۷</td> <td>۱۸</td> <td>۱۷</td> <td>۱۱</td> </tr> </table> الف) در کدام دسته پراکندگی بیشتر است؟ ب) نمرات کدام کلاس بهتر است ؟ چرا ؟	الف	۵	۱۱	۱۹	۱۷	۱۴	۱۳	۱۰	۱۱	۲۰	ب	۹	۷	۱۸	۱۳	۱۹	۱۷	۱۸	۱۷	۱۱	۱۱
الف	۵	۱۱	۱۹	۱۷	۱۴	۱۳	۱۰	۱۱	۲۰													
ب	۹	۷	۱۸	۱۳	۱۹	۱۷	۱۸	۱۷	۱۱													

