

آمار توصیفی

قسمت اول: قاعده عددی

آمار توصیفی

قسمت اول : قاعده عددی

- معیارهای مکانی
- معیارهای پراکندگی

$$\mu \quad \sigma \quad \%$$
$$\bar{x}$$

معیارهای مکانی

- میانگین
- میانه
- مد
- صدکها
- چارکها

مثال: کرایه های آپارتمان

داده های زیر یک نمونه از مقادیر اجاره ماهانه برای آپارتمانهای یک خوابه است. داده ها یک نمونه 70 تایی از آپارتمانها در یک شهر خاص هستند. داده ها به صورت صعودی مرتب شده اند.

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

میانگین

- میانگین مجموعه ای از داده ها معدل مقدار کل داده هاست.
- اگر داده ها از یک نمونه باشند، میانگین با $\bar{X}$ نشان داده می شود.

$$\bar{x} = \frac{\sum x_i}{n}$$

- اگر داده ها از یک جامعه باشند، میانگین با μ نشان داده می شود.

$$\mu = \frac{\sum x_i}{N}$$

مثال: کرایه های آپارتمان

• میانگین

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

میانہ

- میانہ از شاخصهای مکانی است. بیشتر برای گزارش درآمد سالانه و ویژگی مقدار داده ها به کار می رود.
- افزایش مقدار کمی از درآمدهای کلان یا ویژگی مقادیر می توانند میانگین را زیاد کنند.

میانه

- میانه یک مجموعه داده ها زمانی که ارقام به صورت صعودی مرتب می شوند، مقدار وسطی است.
- برای تعداد مشاهدات فرد ، میانه مقدار وسطی است.
- برای تعداد مشاهدات زوج، میانه متوسط دو مقدار وسطی است.

مثال: کرایه های آپارتمان

• میانه

صدک 50 ام = میانه

$$i = (p/100)n = (50/100)70 = 35.5$$

میانگین مقدار داده های 35 ام و 36 ام

$$\text{میانه} = (475 + 475)/2 = 475$$

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

مد

- مد در یک مجموعه داده ها مقداری است که دارای بیشترین فراوانی باشد.
- بیشترین فراوانی می تواند در دو یا بیش از دو مقدار متفاوت باشد.
- اگر داده ها دقیقاً دو مد داشته باشند، داده ها دو مدی هستند.
- اگر داده ها بیش از دو مد داشته باشند، داده ها چند مدی هستند.

مثال: کرایه های آپارتمان

• مد

450 بیشترین فراوانی را دارد (7 مورد)

$$\text{مد} = 450$$

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

صدکها

- یک صدک اطلاعاتی در باره این که داده ها چگونه روی یک فاصله از کوچکترین مقدار تا بزرگترین مقدار پخش می شوند را فراهم می سازد.
- نمرات امتحانات ورودی کالج ها و دانشگاهها اغلب به صورت درصد گزارش می شود.

صدکها

- صدک p ام یک سری از داده ها مقداری است که دست کم p درصد از موارد را در بر می گیرد یا کمتر و دست کم $(100 - p)$ درصد از داده ها یا بیشتر را در بر می گیرد.

– داده ها به ترتیب صعودی مرتب شده اند.

– i محل صدک p ام را محاسبه می کند.

$$i = (p/100)n$$

– اگر i عدد صحیح نباشد، به رقم بالاتر گرد می کنیم. صدک p ام مقداری است در جایگاه i ام.

– اگر i عدد صحیح باشد، صدک p ام میانگین مقادیر در جایگاه i و $i+1$ است.

مثال: کرایه های آپارتمان

• صدک 90 ام

$$i = (p/100)n = (90/100)70 = 63$$

میانگین مقدار داده های 63 و 64 ام:

$$\text{صدک 90 ام} = (580 + 590)/2 = 585$$

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

چارکها

- چارکها، صدکهای معینی هستند.
- چارک اول = صدک 25 ام
- چارک دوم = صدک 50 ام = میانه
- چارک سوم = صدک 75 ام

مثال: کرایه های آپارتمان

• چارک سوم

چارک سوم = صدک 75 ام

$$i = (p/100)n = (75/100)70 = 52.5 = 53$$

$$= 525 \text{ چارک سوم}$$

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

معیارهای پراکندگی

- اغلب بهتر است که معیارهای متغیر (پراکندگی)، را به خوبی معیارهای مکانی در نظر بگیریم.
- برای مثال، ما در انتخاب کارپرداز A یا کارپرداز B ممکن است نه تنها میانگین زمان تحویل داده ها را برای هر یک در نظر بگیریم، بلکه تغییرات زمانی تحویل داده ها را نیز برای هر کدام در نظر بگیریم.

معیارهای پراکندگی

- دامنه تغییرات
- دامنه میان چارکی
- واریانس
- انحراف معیار
- ضریب تغییرات

دامنه تغییرات

- دامنه تغییرات مجموعه داده ها، تفاوت بین بزرگترین و کوچکترین مقدار داده هاست.
- از شاخصهای ابتدائی متغیرهاست.
- نسبت به بزرگترین و کوچکترین مقدار داده خیلی حساس است.

مثال: کرایه های آپارتمان

• دامنه تغییرات

$$\text{کوچکترین مقدار} - \text{بزرگترین مقدار} = \text{دامنه تغییرات}$$
$$= 615 - 425 = 190$$

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

دامنه میان چارکی

- دامنه میان چارکی در یک سری داده ها، تفاوت بین چارک سوم و چارک اول است.
- دامنه تغییرات 50% میانی داده ها است.
- حساسیت را برای مقادیر انتهایی داده ها از بین می برد.

مثال: کرایه های آپارتمان

- دامنه میان چارکی

$$(Q3) = 525 \text{ چارک سوم}$$

$$(Q1) = 445 \text{ چارک اول}$$

$$\text{دامنه میان چارکی} = Q3 - Q1 = 525 - 445 = 80$$

425	430	430	435	435	435	435	435	440	440
440	440	440	445	445	445	445	445	450	450
450	450	450	450	450	460	460	460	465	465
465	470	470	472	475	475	475	480	480	480
480	485	490	490	490	500	500	500	500	510
510	515	525	525	525	535	549	550	570	570
575	575	580	590	600	600	600	600	615	615

واریانس

- میزان تغییراتی که برای تمام داده ها مورد استفاده قرار می گیرد، واریانس نام دارد.
- واریانس بر اساس تفاوت بین مقدار هر مشاهده (X_i) و میانگین ($\bar{X}$) برای یک نمونه و μ برای جامعه است.

واریانس

- واریانس متوسط اختلاف مربعات بین مقدار هر داده و میانگین است.
- اگر مجموعه داده ها یک نمونه باشند، واریانس با s^2 نشان داده می شود.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

- اگر مجموعه داده ها یک جامعه باشند، واریانس با σ^2 نشان داده می شود.

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

انحراف معیار

- انحراف معیار مجموعه داده ها ریشه دوم واریانس است.
- انحراف معیار در همان واحدهای داده اندازه گیری می شود و مقایسه واریانس با میانگین را آسانتر می سازد.
- اگر مجموعه داده ها یک نمونه باشد، انحراف معیار را با S نشان می دهیم.

$$S = \sqrt{S^2}$$

- اگر مجموعه داده ها یک جامعه باشد، انحراف معیار را با σ نشان می دهیم.

$$\sigma = \sqrt{\sigma^2}$$

ضریب تغییرات

- ضریب تغییرات نسبت بزرگی انحراف معیار به میانگین را نشان می دهد.
- اگر مجموعه داده ها یک نمونه باشد، ضریب تغییرات به صورت زیر محاسبه میشود:

$$\frac{s}{\bar{x}}(100)$$

- اگر مجموعه داده ها یک جامعه باشد، ضریب تغییرات به صورت زیر محاسبه می شود:

$$\frac{\sigma}{\mu}(100)$$

مثال: کرایه های آپارتمان

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = 2,996.16$$

- واریانس

$$s = \sqrt{s^2} = \sqrt{2996.47} = 54.74$$

- انحراف معیار

$$\frac{s}{\bar{x}} \times 100 = \frac{54.74}{490.80} \times 100 = 11.15$$

- ضریب تغییرات

پایان، قسمت اول

