

# بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

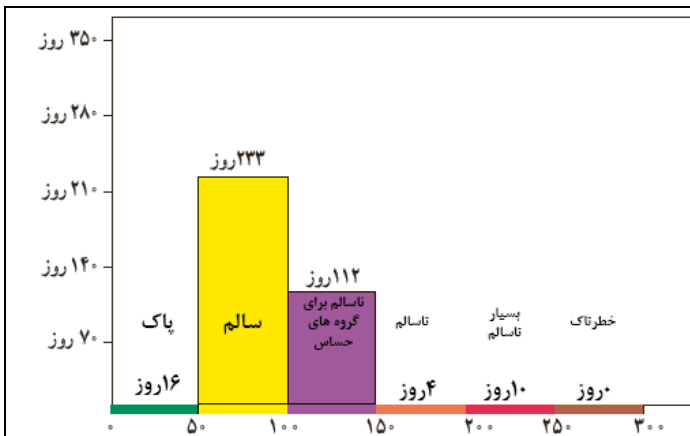
## جزوه فصل ۳ آمار و احتمال یازدهم

درس ۱ (توصیف و نمایش داده ها)

### تعریف ها

- ▶ **داده ها:** واقعیت هایی درباره یک شیء یا فردند که در محاسبه، برنامه ریزی و پیش بینی به کار می روند.
  - ▶ **متغیر:** هر ویژگی از اشیا یا اشخاص، که در اعضای جامعه یکسان نیست و معمولاً از یک عضو به عضو دیگر تغییر می کند را **متغیر** می گویند و عددی که به آن ویژگی یک عضو نسبت داده می شود را **مقدار متغیر**، یا **مشاهده** می گویند.
  - ▶ **فراوانی یک داده:** تعداد دفعاتی که هر داده مشاهده می شود را **فراوانی** آن داده می گویند.
  - ▶ **فراوانی نسبی یک داده:** با تقسیم فراوانی هر داده به تعداد کل داده ها، **فراوانی نسبی** آن داده به دست می آید.
- نکته: اگر فراوانی نسبی داده ها در ۱۰۰ ضرب شود، آن گاه **درصد داده ها** به دست می آید.

### مثال و پاسخ ؟



نسبت به نمودار روبرو به سوالات پاسخ دهید.  
چند درصد از روزهای سال، هوا سالم بوده است؟

$$\frac{233}{350} \times 100 = 62\%$$

چند درصد روزهای سال، هوا ناسالم و بسیار ناسالم بوده است؟

$$\frac{4 + 10}{350} \times 100 = 4\%$$

کدام نمودار، در پاسخ دادن به سوالات، ما را بهتر راهنمایی می کند؟ **نمودار فراوانی نسبی**

- برای متغیرهای پیوسته از نمودار بافت نگار استفاده می شود.
- برای متغیرهای گسسته از نمودارهای دایره ای و میله ای استفاده می شود.
- برای متغیرهای کیفی از نمودارهای دایره ای و میله ای استفاده می شود.

### درس دوم (معیار های گرایش به مرکز)

**مجموع داده ها:** اگر  $n$  داده  $x_1, x_2, \dots, x_n$  داشته باشیم. مجموع داده ها را با نماد سیگما ( $\Sigma$ ) نمایش می دهیم.

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$$

■ عبارت بالا یا  $\sum_{i=1}^n x_i$  می خوانیم سیگمای  $n$  از  $1$  تا  $n$ ،  $x_i$  می خوانیم.

**میانگین یا متوسط داده ها:** میانگین یا متوسط داده ها را با نماد  $(\bar{x})$  نمایش داده می شود. و به صورت زیر تعریف می شود.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

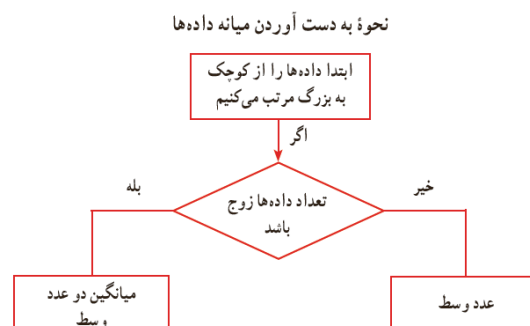
**میانگین موزون داده ها:** اگر  $n$  داده  $x_1, x_2, \dots, x_n$  داشته باشیم به طوری که هر یک از این داده ها دارای تعداد تکرار  $w_1, w_2, \dots, w_n$  هستند. که به هر یک از آنها وزن داده متناظر با آن می گوئیم. میانگین موزون داده ها را داده ها را با نماد  $\bar{x}_w$  نمایش می دهیم و به صورت زیر تعریف می کنیم.

$$\bar{x}_w = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i} = \frac{x_1 w_1 + x_2 w_2 + \dots + x_n w_n}{w_1 + w_2 + \dots + w_n}$$

**میانه:** عدد وسط مجموعه ای از داده ها را که از کوچک به بزرگ مرتب شده باشند میانه داده ها می گوئیم و آن را با  $Q_2$  نشان می دهیم.

**چارک اول:** میانه یک چهارم اول داده های مرتب شده را چارک اول داده ها گوئیم و آن را با  $Q_1$  نشان می دهیم.

**میانه سه چهارم داده های مرتب شده را چارک سوم گوئیم و آن را با  $Q_3$  نشان می دهیم.**



سوال: میانه، چارک اول و سوم داده های جمع آوری شده زیر را مشخص نمایید.

۱۶-۴-۳۰-۱۸-۱۲-۲۶

جواب: ابتدا داده ها را از کوچک به بزرگ مرتب می کنیم

۴-۱۲-۱۶-۱۸-۲۶-۳۰

$$\frac{26+30}{2} = 27 \text{ چارک سوم}$$

$$\frac{4+12}{2} = 8 \text{ چارک اول}$$

$$\frac{16+18}{2} = 17 \text{ میانه}$$

▶ **مد یا نما داده ها:** داده ای که بیشترین فراوانی را داشته باشد، مد یا نما داده ها نام دارد.

■ نکته: اگر در داده ها، همه داده ها یک فراوانی داشته باشند، آن گاه این داده ها مد ندارند. و اگر در داده هایی، دو

داده بیشترین فراوانی را داشته باشند، آن گاه این داده ها **دو مد** دارند.

### مثال و پاسخ ②

سوال: مد داده های جمع آوری شده زیر را مشخص نمایید.

۷ و ۲ و ۳ و ۹ و ۸ و ۸

جواب: ابتدا فراوانی هر عدد را می نویسیم یا نمودار بافت نگار این اعداد را رسم می کنیم.

۲	۱ عدد	۷	۱ عدد
۳	۱ عدد	۸	۲ عدد
۶	۱ عدد	۹	۱ عدد

عدد ۸ در این داده ها مد است زیرا ۲ عدد موجود است.

▶ **داده دورافتاده:** مشاهده ای (متغیری) که تفاوت بسیار زیادی با سایر مشاهدات مجموعه داده ها داشته باشد،

میانگین داده ها را تحت تأثیر قرار داده در حالی که تأثیری بر میانه و مد داده ها ندارد.

■ نکته: در تفسیر و تحلیل مسائل آماری، در نظر گرفتن تنها یک شاخص گرایش به مرکز **کافی نیست** می بایست هر

سه معیار **میانگین، میانه و مد** محاسبه شود و بر اساس هدف مورد بررسی، معیار مناسب انتخاب و از آن برای

انجام تفسیر، **قضای و پیش بینی** مورد استفاده قرار گیرد.

### مثال و پاسخ ②

سوال: میانگین داده های زیر را حساب کنید.

۱۰ و ۳ و ۳ و ۴ و ۲

جواب: برای محاسبه جواب ۱۰ را در نظر نمی گیریم.

$$\frac{4 + 3 + 3 + 2}{4} = 3$$

▶ **انحراف از میانگین:** اختلاف هر کدام از داده ها با میانگین داده ها انحراف از میانگین می گویند.  
 ▶ اگر مجموع انحراف از میانگین ها برابر با صفر خواهد شد و این به دلیل آن است که برخی از داده ها از میانگین بزرگ تر و برخی دیگر کوچک ترند در نتیجه مقادیر مثبت و منفی حاصل می شوند که مجموع آنها همدیگر را خنثی می کنند.

■ **راحل:** توان دوم انحراف از میانگین داده ها در نظر گرفته می شود.

▶ **انحراف معیار:** در آمار، یک معیار سنجش برای میزان پراکندگی داده ها حول میانگینشان، **انحراف معیار** است.

▶ **محاسبه انحراف معیار:** اگر  $n$  داده از جامعه به صورت  $x_1, x_2, \dots, x_n$  داشته باشیم و با نماد  $\sigma$  نمایش می دهیم.

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

■ نکته: در رابطه بالا  $x_1 - \bar{x}$  را انحراف  $i$ ام از میانگین داده ها می گویند.

▶ **واریانس داده ها:** توان دوم انحراف معیار داده ها را واریانس داده ها گویند و آن را با نماد  $\sigma^2$  نمایش می دهیم.

### نکات انحراف معیار

■ اگر انحراف معیار مجموعه داده ها عدد کوچکی باشد (کمتر از میانگین باشد)، بدین معناست که پراکندگی داده ها حول میانگینشان کم و در نتیجه داده ها به هم نزدیک تر است.

■ اگر انحراف معیار مجموعه داده ها عددی بزرگ باشد (بیشتر از میانگین باشد)، بدین معناست که پراکندگی داده ها حول میانگینشان زیاد و در نتیجه داده ها از هم دورتر است.

▶ ضریب تغییرات داده ها: معیاری است که از تقسیم انحراف معیار داده ها  $\sigma$  به میانگین داده ها  $\bar{x}$  به دست می آید و آن را با نماد CV نشان می دهیم.

$$CV = \frac{\sigma}{\bar{x}}$$

■ نکته: هر قدر ضریب تغییرات کمتر باشد، میزان پراکندگی داده ها کمتر خواهد شد که این موضوع برای ما مطلوب است. یعنی: **ضریب تغییرات  $\propto$  میزان پراکندگی داده ها**

**نمودار جعبه ای:** در این نمودار چارک اول، میانه، چارک سوم، بیشترین و کمترین مقدار داده ها به طور هم زمان نشان داده می شود.

