بسمه تعالی

موضوع : مبانی آمار و احتمال برای ورود به علم داده و تحلیل داده

بخش هفتم:

- **معیارهای پراکندگی:**
  - دامنه تغییرات(Range)
  - واریانس(Variance)
  - انحراف معیار(Standard Deviation)

در این پارت از اموزش به بررسی **معیارهای پراکندگی داده ها** می پردازیم .

در اموزش های قبلی درباره آمار **معیار تمرکز صحبت** کردیم .

"معیارهای مرکزی به ما می‌گویند داده‌ها کجا قرار گرفته‌اند، اما معیارهای پراکندگی به ما می‌گویند داده‌ها چقدر از هم فاصله دارند".

معیارهای پراکندگی چیست؟

معیارهای پراکندگی شاخص‌هایی هستند که میزان پراکندگی یا فاصله داده‌ها را نسبت به یکدیگر یا نسبت به میانگین نشان می‌دهند.

هرچه داده‌ها به هم نزدیک‌تر باشند، پراکندگی کمتر است . هرچه فاصله آن‌ها بیشتر باشد، پراکندگی نیز بیشتر خواهد بود. سه آماره مهم پراکندگی عبارت‌اند از:

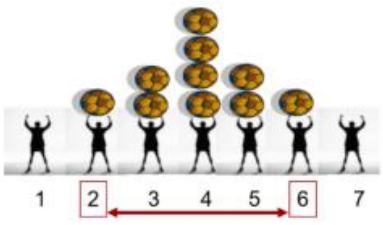
- دامنه تغییرات(Range)
- واریانس(Variance)
- انحراف معیار (Standard Deviation)

دامنه تغییرات (Range):

ساده‌ترین معیار پراکندگی، دامنه تغییرات است.

مثال : ۱ و ۲و ۳و ۴و ۵و ۶و ۷

دامنه: ۶

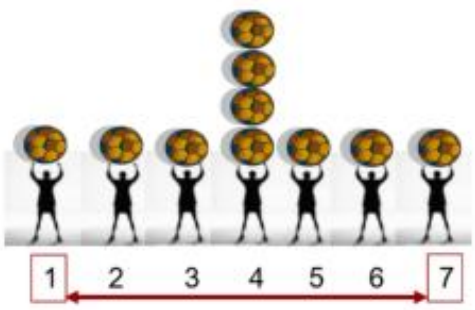


$R = X_{max} - X_{min}$: دامنه

واریانس(Variance) :

فرض کنید بخواهیم بدانیم داده‌ها به طور متوسط چقدر از میانگین فاصله دارند. اگر فقط اختلاف هر داده با میانگین را جمع کنیم، نتیجه همیشه صفر می‌شود. زیرا اختلاف‌های مثبت و منفی یکدیگر را خنثی می‌کنند. برای حل این مشکل، اختلاف‌ها را به توان دو می‌رسانیم. سپس میانگین این مربع اختلاف‌ها را محاسبه می‌کنیم. به این مقدار واریانس گفته می‌شود.

مثال : ۱ و ۲و ۳و ۴و ۵و ۶و ۷



میانگین: ۴

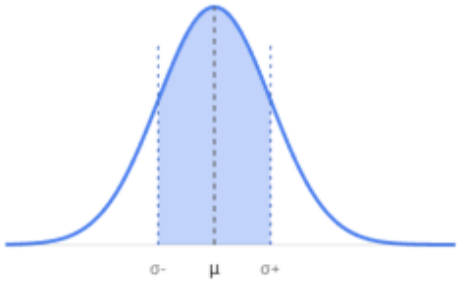
واریانس : ۴.۷

واریانس : $S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$

انحراف معیار(Standard Deviation) :

انحراف معیار در واقع ریشه دوم واریانس است. به همین دلیل، واحد آن دوباره به واحد اصلی داده‌ها برمی‌گردد.

مثال : ۱ و ۲و ۳و ۴و ۵و ۶و ۷



انحراف معیار: ۲.۲

انحراف معیار : $S = \sqrt{S^2}$