بسم الله الرحمن الرحیم

دانشگاه فنی و حرفه ای دختران ارومیه

درس آمارواحتمالات

جلسه ششم

استاد: اکرم سلطان پور

پارامترهای پراکندگی

این پارامترها کمیت‌هایی هستند که چگونگی تغییرات داده‌ها را نشان می‌دهند و انواع مختلف دارند

۱) **دامنه تغییرات:** دامنه تغییرات که آن را با نماد R نشان می‌دهیم، عبارت است از تفاضل بین بزرگ‌ترین و کوچک‌ترین داده.

$$R = \text{بزرگترین داده} - \text{کوچک‌ترین داده}$$

ملاحظه می‌شود که محاسبه R ساده است ولی با توجه به این‌که فقط به دو داده بزرگ و کوچک بستگی دارد، بنابراین سایر داده‌ها در این پارامتر نقشی ندارند. لذا این پارامتر در مواردی می‌تواند گمراه کننده باشد و بنابراین زیاد مورد استفاده قرار نمی‌گیرد.

بدیهی است که هر چقدر R بزرگ باشد، دلیل بر پراکنده‌تر بودن داده‌های مربوطه می‌باشد.

۱–۲۹. مثال: نمرات دو دانش‌آموز در جدول زیر درج شده است.

	زبان فارسی	شیمی	ریاضی	جغرافیا
دانش آموز اول	۱۳	۱۳	۱۲	۱۴
دانش آموز دوم	۱۸	۱۳	۸	۱۳

میانگین و دامنه تغییرات نمرات این دو دانش آموز را محاسبه کنید.

حل:

$$\mu_۱ = \frac{\sum x_i}{n} = \frac{۵۲}{۴} = ۱۳ \quad ; \quad R_۱ = ۱۴ - ۱۲ = ۲$$

$$\mu_۲ = \frac{\sum x_i}{n} = \frac{۵۲}{۴} = ۱۳ \quad ; \quad R_۲ = ۱۸ - ۸ = ۱۰$$

بنابراین با این که هر دو دانش آموز دارای میانگین نمرات یکسانی هستند ولی دانش آموز دوم دارای پراکندگی نمرات بیشتری است.

۲) انحراف متوسط: اختلاف بین داده‌های یک مجموعه اطلاعات آماری از میانگین را انحراف می‌نامند. انحراف نشان می‌دهد که یک مشاهده تا چه میزان با میانگین متفاوت است.

$$\text{انحراف} = x - \mu$$

برای مثال، تصور کنید زمان تحویل کالایی که در یک هفته خاص به کارخانه آلفا سفارش شده است، از زمان درخواست، به گونه‌ای باشد که در جدول زیر آمده است.

زمان تحویل به روز x	انحراف $x - \mu$	قدر مطلق انحراف $\lvert x - \mu \rvert$
۸	۱	۱
۹	۲	۲
۶	۱-	۱
۴	۳-	۳
۸	۱	۱
$\sum x = ۳۵$	$\sum (x - \mu) = ۰$	$\sum \lvert x - \mu \rvert = ۸$

یعنی از زمان درخواست به طور متوسط هفت روز طول می‌کشد تا کالای درخواستی تحویل داده شود. اما سؤال دیگری که مطرح است آن است که مدت زمان تحویل کالا تا چه اندازه متغیر است. مسلماً هرچه اختلاف بین زمان تحویل کالا از یک سفارش به سفارش دیگر کمتر باشد، به عبارت دیگر، هر چه پراکندگی مدت زمان تحویل کالا کمتر باشد، امر برنامه ریزی موجودی انبار راحت تر است. برای ساختن شاخصی که پراکندگی را نشان می‌دهد، شاید بهتر است میانگین انحرافات داده‌ها را

محاسبه کنیم. هر چقدر این میانگین بزرگتر باشد، بیان‌گر آن است که اختلاف داده‌ها از میانگینشان بیشتر و در نتیجه پراکندگی داده‌ها بیشتر است. ولی متأسفانه چون جمع جبری انحرافات از میانگین $\left(\sum(x-\mu)\right)$ همواره صفر است، میانگین انحرافات چیزی را به دست نمی‌دهد. اما اگر بتوان به گونه‌ای از قید علامت‌های منفی انحرافات از میانگین خلاص شد، می‌توان متوسط انحرافات از میانگین را محاسبه کرد. یکی از راه حل‌ها این است که قدر مطلق انحرافات را در نظر بگیریم و متوسط قدر مطلق انحرافات را به عنوان شاخص پراکندگی مورد استفاده قرار دهیم. این شاخص پراکندگی را اصطلاحاً متوسط قدر مطلق انحرافات یا به طور خلاصه انحراف متوسط $(A.D)$ می‌نامند.

$$A.D=\frac{\sum|x-\mu|}{N}$$

و اگر داده‌ها طبقه‌بندی شده و دارای فراوانی باشند. داریم:

$$A.D=\frac{\sum|x_i-\mu|f_i}{\sum f_i}$$

بنابراین در مثال فوق، داریم:

$$A.D=\frac{\sum|x-\mu|}{N}=\frac{۸}{۵}=۱/۶$$

بنابراین می‌توان گفت مدت زمان تحویل یک سفارش خاص ممکن است به طور متوسط ۱/۶ روز بیشتر و یا کمتر از میانگین باشد.

۳) **واریانس و انحراف معیار:** همان‌طور که در قسمت قبل بیان شد، برای رهایی از علامت‌های منفی انحرافات از میانگین، قدر مطلق انحرافات را در نظر گرفتیم. راه دیگر آن است که انحراف را به توان ۲ برسانیم و به این ترتیب از قید علامت‌های منفی خلاص شویم. اکنون با محاسبه متوسط این مجذور انحرافات، شاخص پراکندگی دیگری به دست می‌آید که به آن واریانس یا پراش می‌گوییم و با علامت σ^2 یا S^2 نشان می‌دهیم. $\sigma^2 = \frac{\sum(x-\mu)^2}{N}$ می‌باشد و اگر داده‌ها در جدول طبقه‌بندی شده باشند، داریم: $\sigma^2 = \frac{\sum(x-\mu)^2 f_i}{\sum f_i}$.

با توجه به مثال فوق، جدول زیر را تشکیل می‌دهیم.

زمان تحویل به روز x	انحراف $x-\mu$	مجذور انحرافات $(x-\mu)^2$
۸	۱	۱
۹	۲	۴
۶	۱-	۱
۴	۳-	۹
۸	۱	۱
$\sum x = ۳۵$	$\sum(x-\mu) = ۰$	$\sum(x-\mu)^2 = ۱۶$

واریانس مدت زمان تحویل کالا با توجه به جدول فوق، به صورت زیر محاسبه می‌شود.

$$\sigma^{۲} = \frac{\sum (x-\mu)^{۲}}{N} = \frac{۱۶}{۵} = ۳/۲$$

عیب واریانس آن است که واحدی ندارد. مثلاً اگر داده‌ها بر حسب متر باشد، آنگاه $(x-\mu)$ نیز بر حسب متر است ولی $(x-\mu)^{۲}$ بر حسب متر مربع خواهد بود و لذا قابل مقایسه نخواهد بود. همچنین، عمل توان‌رسانی نه تنها مقیاس اندازه‌گیری واریانس را بی مفهوم می‌کند، بلکه انحرافات را بزرگ می‌کند. برای خنثی کردن این اثر می‌توان جذر مثبت واریانس را محاسبه کرد. به این ترتیب شاخص پراکندگی دیگری به دست می‌آید که به انحراف استاندارد یا انحراف معیار معروف است و با F و σ نشان می‌دهیم که برابر است با:

$$\sigma = \sqrt{\frac{\sum (x-\mu)^{۲}}{N}}$$

به عنوان مثال انحراف معیار مدت زمان تحویل کالا، با توجه به جدول فوق، برابر است با:

$$\sigma = \sqrt{\frac{۱۶}{۵}} = \sqrt{۳/۲} = ۱/۷۹$$

۱–۳۱. مثال: مدت زمانی که دانشجویان کلاس، صرف مطالعه جهت آمادگی امتحان میان ترم کرده‌اند، در جدول زیر ثبت شده است. انحراف متوسط و انحراف معیار مدت زمان مطالعه را محاسبه کنید.

زمان مطالعه به ساعت	تعداد دانشجویان f_i
۰–۴	۳
۴–۸	۸
۸–۱۲	۷
۱۲–۱۶	۵
۱۶–۲۰	۲
	$N=$ ۲۵

حل:

x_i	$x_i f_i$	$x_i-\mu$	$\lvert x_i-\mu\rvert$	$(x_i-\mu)^۲$
۲	۶	۷/۲−	۷/۲	۵۱/۸۴
۶	۴۸	۳/۲−	۳/۲	۱۰/۲۴
۱۰	۷۰	۰/۸	۰/۸	۰/۶۴
۱۴	۷۰	۴/۸	۴/۸	۲۳/۰۴
۱۸	۳۶	۸/۸	۸/۸	۷۷/۴۴
	$\sum x_i f_i$		۲۴/۸	۱۶۳/۲

$$\mu = \frac{\sum x_i f_i}{N} = \frac{۲۳۰}{۲۵} = ۹/۲$$

$$A.D = \frac{\sum |x_i - \mu| f_i}{N} = \frac{(۷/۲ \times ۳) + (۳/۲ \times ۸) + \ldots + (۸/۸ \times ۲)}{۲۵} = \frac{۹۴/۴}{۲۵} \cong ۳/۷۸$$

$$\sigma^۲ = \frac{\sum (x_i - \mu)^۲ f_i}{N} = \frac{(۵۱/۸۴ \times ۳) + (۱۰/۲۷ \times ۸) + \ldots + (۷۷/۴۴ \times ۲)}{۲۵} = \frac{۵۱۲}{۲۵} = ۲۰/۴۸$$

$$\sigma = \sqrt{۲۰/۴۸} \cong ۴/۵۳$$

با استفاده از خواص $\sum$، به عنوان فرمول ساده‌تر، می‌توان واریانس را از رابطه زیر نیز محاسبه کرد.

$$\sigma^۲ = \mu_{x^۲} - (\mu_x)^۲$$

که در آن $\mu_{x^۲}$ میانگین توان دوم داده‌ها و μ_x میانگین داده‌ها است. به عبارت دیگر:

$$\sigma^۲ = \frac{\sum x_i^۲ f_i}{N} - \left(\frac{\sum x_i f_i}{N}\right)^۲$$

تمرین

در جامعه ایی توزیع فراوانی صفت متغییر به صورت زیر بدست آمده است

x	2	4	6	8	10
Fi	10	15	22	18	5

مطلوبست محاسبه

الف–میانگین ب–واریانس ج–انحراف معیار د– انحراف متوسط

موفق و سلامت باشید