آمار و احتمال

پایه ی یازدهم «رشته ی ریاضی فیزیک»

فصل ۳ : آمار توصیفی

تهیه کننده : جابرعامری

دبیر ریاضی شهرستان های اهواز و باوی

www.mathtower.ir

@amerimath

مهر ۱۴۰۰

درس اوّل : توصیف و نمایش داده ها

همانطور که می دانید، در یک فعالیت آماری بعد از گردآوری داده ها، تنظیم و رده بندی و خلاصه کردن آنها، جهت نتیجه گیری و تعیین شاخص ها ، لازم است. در این درس به طبقه بندی داده ها و رسم نمودار های آماری می پردازیم.

دسته بندی اطلاعات

می دانیم که واقعیت هایی درباره های یک شیء یا فرد که در محاسبه ، برنامه ریزی و پیش بینی به کار می روند را داده (اطلاعات) می نامند. واضح است که باید داده های جمع آوری شده در هر زمینه ای که باشند را دسته بندی نمود تا از این طریق بتوان به راحتی داده ها را توصیف و ازآنها نتیجه‌گیری کرد، یکی از روش های دسته بندی داده ها تشکیل **جدول فراوانی** موسوم به توزیع فراوانی است در واقع یکی از کارآمدترین روشها برای خلاصه کردن و سازمان بندی کردن اطلاعات توزیع فراوانی می باشد.که در اینجا به توضیح آن می‌پردازیم.

جدول فراوانی

در یک مجموعه ی داده های آماری به تعداد دفعاتی که هر داده تکرار می شود، **فراوانی مطلق** یا به اختصار فراوانی آن داده گویند.

مثال: نمرات مسئولیت پذیری بیست مدیر به شرح زیراست، فراوانی این داده‌ها را مشخص کنید.

x_i : ۵ و۱۷و۱۲و۵ و۱۲و۱۲و۵ و ۶ و۱۲و۶ و۱۰و۵ و۱۵و۱۵و۱۲و۶ و۱۵و۱۰ و۱۵و۱۲

$F(x=۱۲)=۶$ فراوانی۱۲برابر ۶ است.

$F(x=۵)=۴$ فراوانی ۵ برابر ۴ است.

$F(x=۱۵)=۴$ فراوانی۱۵برابر ۴ است.

$F(x=۶)=۳$ فراوانی ۶ برابر ۳ است.

$F(x=۱۷)=۱$ فراوانی۱۷برابر ۱ است.

$F(x=۱۰)=۲$ فراوانی۱۰برابر ۲ است.

وبه طور خلاصه داریم.

نمره (x_i)	۵	۶	۱۰	۱۲	۱۵	۱۷	جمع
فراوانی (F_i)	۴	۳	۲	۶	۴	۱	۲۰

نتیجه: واضح است که جمع فراوانی های مطلق با اندازه نمونه یا جامعه برابر است.

روشهای دسته بندی داده ها

هر ویژگی از اشیاء یا اشخاص که در اعضای جامعه یکسان نیست و معمولاً از یک عضو به عضو دیگر تغییر می کند را متغیر می نامند. نماد یا عددی که به آن ویژگی یک عضو نسبت داده می شود را مقدار متغیر (مشاهده) می گویند. متغیر ها براساس اینکه مقادیر عددی می گیرند یا خیر، دارای دو نوع اصلی کمی و کیفی می باشد،

متغیر کمی : متغیرهایی هستند که مقادیر عددی می گیرند و برای آنها عملیات ریاضی از قبیل جمع ، تفریق و معدل گیری قابل انجام است.

مثال : سن ، نمره ، تعداد افراد خانواده ، طول مکالمه‌ی تلفنی و ....

متغیر کیفی : متغیر هایی هستند که صرفاً برای دسته بندی افراد یا اشیاء در گروه ها به کار می روند و لزوماً مقدار عددی نمی گیرند.

مثال : جنسیت (مرد و زن) ، کیفیت کالا (مرغوب، متوسط ، نامرغوب)

هدف علم آمار نتیجه گیری از تعداد زیادی داده است که از جامعه یا نمونه بدست می‌آیند. تنظیم جدول آماری جهت توصیف نتایج مفید است. دسته بندی داده‌ها با توجه به نوع متغیر آنها به صورت های مختلف انجام می شود.

الف) دسته بندی کیفی: این روش برای دسته بندی داده‌هایی استفاده می‌شود که دارای متغیر کیفی می باشند.

مثال: گروه خونی ۲۵ دانش آموز یک کلاس به شرح زیر است.

A,O,AB,O,B,A,O,B,B,O,B,AB,A,AB,O,A,A,B,A,B,B,O,O,AB,O

حال می توان این مشاهدات را به صورت زیر دسته بندی کرد.

ردیف	دسته (گروه خونی)	خط نشان (T)	فراوانی(F)
۱	A	$\cancel{||||}$ /	۶
۲	B	$\cancel{||||}$ //	۷
۳	AB	////	۴
۴	O	$\cancel{||||}$ ///	۸
جمع	–	–	۲۵

ب) دسته بندی کمی: این روش برای دسته بندی داده هایی استفاده می‌شود کـه دارای متغیرکمـی مـی باشند.

مثال: موضوع یک تحقیق آماری بررسی سطح نمرات درس ریاضی دانش آموزان یـک مدرسـه اسـت،برای انجام این تحقیق مقرر شد یک نمونه‌ی ۳۶ نفری به صورت تصادفی انتخاب و نمرات دانش آموزان انتخـاب شده از دفتر مدرسه دریافت شود، پس ازمراجعه به دفتر نمرات زیر بدست آمد.

۱۲/۵و۱۳و۱۲و۱۳/۵و۱۴و۱۱و۱۳و۹و۸ و۱۵و۱۴و۱۹و۱۵و۱۴و۱۰/۵و۱۱/۵و۱۵و۱۸و۱۷ ۱۴و۱۶و۱۸و۱۵و۹/۵و۶

۱۲/۲۵و۲۰و۱۸/۵و۱۷و ۱۵/۵و۱۶/۷۵و۱۱و۸/۵ و۷ و۳ و۱۳

حال می توان این نمرات را به صورت زیر دسته بندی کرد.

ردیف	دسته ها (CL)		خط نشان (T)	فراوانی مطلق (F_i)
۱	نمرات از صفر تا ۴	[۰و۴)	/	۱
۲	نمرات از ۴ تا ۸	[۴و ۸)	//	۲
۳	نمرات از ۸ تا ۱۲	[۸و۱۲)	~~////~~ ///	۸
۴	نمرات از ۱۲ تا ۱۶	[۱۲و۱۶)	~~////~~ ~~////~~ ~~////~~ /	۱۶
۵	نمرات از۱۶تا و با ۲۰	[۱۶و۲۰]	~~////~~ ////	۹
جمع	–	–	–	۳۶

تذکر:

۱: به اعدادی که در دو طرف یک دسته قرار دارند، حدود آن دسته می‌گویند. مثلاًدر دسته‌ی سوّم جدول فوق عدد۸ حد پایین وعدد ۱۲حد بالای آن است.

۲ : سعی می شود که در یک جدول فراوانی، بجز دسته‌ی آخر حد بالای هر دسته جزء آن دسته نباشد. یعنی حدود هر دسته به شکل $[a,b)$ ولی دسته آخر به شکل $[a,b]$ است. گرچه می توان این قـرار داد را تغییـر داد ولی باید در جدول فراوانی قرار داد جدید به خوبی معلوم شود.

۳ : تفاضل حد بالا(یا حد پایین) هر دو دسته متوالی عدد ثابتی است که آن را طول یا فاصله‌ی دسته ها مـی نامند و با C نمایش می دهند. مثلاً : $C = ۱۶ - ۱۲ = ۴$

۴ : در هرجدول فراوانی تعداد داده هایی که در یک دسته قرار می گیرند، فراوانی مطلق آن دسته می گویند.

۵ : در هر جدول فراوانی مجموع فراوانی های مطلق تمام دسته ها با اندازه‌ی نمونه یا جامعه برابر است.

تمرین برای حل :

۱: داده های آماری زیر را در نظر بگیرید.

۲۶و۲۵و۲۴و۲۳و۲۲و۲۱و۲۱و۲۰و۱۹و۱۹و۱۸و۱۷و۱۶و۱۵و۱۴و۱۴و۱۳و۱۲و۱۱و۱۱و۱۰و۱۰و۱۰و۲۹و۲۸و۲۷

جدول فراوانی برای این داده ها تنظیم نموده و طول دسته ها را ۵ بگیرید.

۲ : تعداد افراد خانواده های یک شهرک در یک نمونه۴۰ تایی به صورت زیر است.

۸ و۷ و۳ و۶ و۵ و۴و۲و۴ و۳و۳و۱و۶ و۲و۴ و۵ و ۵ و۴و۲و۳

۷و ۵ و۳ و۴ و۱و۳ و۳ و۴ و۴ و۲ و۶ و۲و۲و۳و۵ و۱و۵ و۲و۱و۲و۴و۳

جدول فراوانی برای این داده ها بر حسب تعداد افراد خانوار تنظیم کنید.

***

معرفی چند مفهوم در جدول آماری

در یک جدول فراوانی معمولاً با مفاهیم زیر روبرو می شویم، که در زیر تعریف آنها را بیان می کنیم.

۱ : مرکز دسته ها (x_i) در یک جدول آماری میانگین حدود هر دسته را مرکزآن دسته می نامند.

$$x_i = \frac{\text{حد بالای آن دسته} + \text{حد پایین هر دسته}}{۲}$$

۲ : فراوانی نسبی(f_i) در یک جدول آماری نسبت فراوانی مطلق هر دسته بر تعداد داده ها را فراوانی نسبی آن دسته می نامند.

$$f_i = \frac{\text{فراوانی مطلق هر دسته}}{\text{تعداد داده ها}}$$

۳ : درصد فراوانی نسبی (P_i) در یک جدول آماری حاصل ضرب فراوانی نسبی هر دسته در عدد۱۰۰را درصد فراوانی نسبی آن دسته می نامند.

$$P_i = f_i \times ۱۰۰$$

۴ : زاویه‌ی قطاع (α_i) در یک جدول آماری حاصل ضرب فراوانی نسبی هر دسته درعدد۳۶۰ را زاویه‌ی قطاع (زاویه‌ی مرکزی متناظر) آن دسته می نامند.

$$\alpha_i = f_i \times ۳۶۰$$

مثال: جدول داده شده را کامل کنید.

مرکز دسته ها	زاویه‌ی قطاع	درصد فراوانی نسبی	فراوانی نسبی	فراوانی مطلق	دسته ها	ردیف
۷	۴۳/۲	۱۲	۰/۱۲	۳	۵ – ۹	۱
۱۱	۷۲	۲۰	۰/۲۰	۵	۹ – ۱۳	۲
۱۵	۱۰۰/۸	۲۸	۰/۲۸	۷	۱۳ – ۱۷	۳
۱۹	۸۶/۴	۲۴	۰/۲۴	۶	۱۷ – ۲۱	۴
۲۳	۴۳/۲	۱۲	۰/۱۲	۳	۲۱ – ۲۵	۵
۲۷	۱۴/۴	۴	۰/۰۴	۱	۲۵ – ۲۹	۶
–	۳۶۰	۱۰۰	۱	۲۵	–	جمع

تمرین۳ : جدول زیر را کامل کنید.

مرکز دسته ها	زاویه‌ی قطاع	درصد فراوانی نسبی	فراوانی نسبی	فراوانی مطلق	دسته ها	ردیف
				۳	۱ – ۵	۱
				۷	۵ – ۹	۲
				۶	۹ – ۱۳	۳
				۴	۱۳ – ۱۷	۴
					–	جمع

نتیجه :

۱ : تفاضل مراکز دو دسته متوالی با طول دسته‌ها برابر است. $(C = x_{i+۱} - x_i)$

۲ : مجموع فراوانی های نسبی تمام دسته ها برابر یک است.

$$(f_۱ + f_۲ + f_۳ + \ldots + f_n = \sum f_i = ۱)$$

۳ : مجموع درصد فراوانی های نسبی تمام دسته ها برابر ۱۰۰ است.

$$(P_۱ + P_۲ + P_۳ + \ldots + P_n = \sum P_i = ۱۰۰)$$

۴ : مجموع زاویه های قطاع های تمام دسته ها برابر ۳۶۰ است.

$$(\alpha_۱ + \alpha_۲ + \alpha_۳ + \ldots + \alpha_n = \sum \alpha_i = ۳۶۰)$$

تمرین برای حل :

۴ : در جدول زیر جای خالی را کامل کنید.

دسته ها	فراوانی مطلق	فراوانی نسبی	مرکز دسته ها	زاویه‌ی قطاع
۲ –۵	۲	۰/۲	۳/۵	
۵ –۸			۶/۵	۱۰۸
	۵	۰/۵		۱۸۰
جمع	۱۰		–	

۵ : داده های زیر یک نمونه‌ی ۲۵تایی از نمرات دانش آموزان پایه‌ی دهم یک منطقه، در مسابقه‌ی علمی ۳۰ نمره ای می باشند.

۱۸و۱۳/۵و۱۴/۷۵و۱۸و۱۴و۱۲/۵و۱۶/۷۵و۱۶و۱۷/۵و۲۱/۵و۲۱و۲۲/۵

۱۶/۵و۱۵/۵و۱۵و۲۰/۵و۲۱و۱۹/۷۵و۱۹و۱۳/۵و۱۴و۱۷و۱۶و۱۸و۲۰

جدول فراوانی زیر را برای این داده ها ، تکمیل کنید.

دسته ها	فراوانی مطلق	فراوانی نسبی	درصد فراوانی نسبی	زاویه‌ی مرکزی
جمع				

۶ : از ۲۵ دانش آموز پایه‌ی یازدهم پیرامون رنگ مورد علاقه برای خودرو پرسیده شد و جدول فراوانی زیر تنظیم گردید. این جدول را تکمیل کنید.

دسته ها	فراوانی مطلق	فراوانی نسبی	درصد فراوانی نسبی	زاویه‌ی مرکزی
سفید	۶			
نقره ای	۸			
یشمی	۷			
آبی	۴			
جمع	۲۵			

۷ : در یک جدول فراوانی، اگر فراوانی نسبی مربوط به گروه خونی B برابر ۰/۴ باشد و مجموع فراوانی های همه‌ی گروه ها ی خونی برابر ۲۰ باشد. فراوانی گروه خونی B را تعیین کنید.

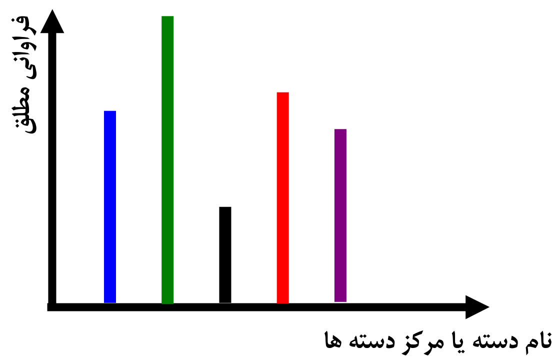
نمودار های آماری

نمایش هندسی داده های دسته بندی شده را نمودار (شاخص هندسی) می‌گویند.نمودارها وسیله ای سودمند برای به تصویر درآوردن و تجسم جامعه یا نمونه می باشند، بطوری که تفهیم نتایج به کمک آنها درکمترین زمان و باساده ترین بیان صورت می گیرد.

نمودار های آماری با توجه به نوع متغیر های بکار رفته در داده ها دارای سه دسته اصلی زیر می باشند.به جدول زیر دقت شود.

نوع متغیرهای بکار رفته در داده های آماری	نام نمودارهای مناسب
کیفی	نمودار میله ای ، نمودار دایره ای
کمی گسسته	نمودار میله ای ، نمودار دایره ای
کمی پیوسته	نمودار بافت نگاشت(مستطیلی) ، نمودار چند ضلعی

نمودارمیله‌ای (ستونی): در این نمودار مهم این است که گروه های متفاوت جامعه یا نمونه از نظر فراوانی مقایسه شوند. محور افقی این نمودار شامل دسته‌ها (نام متغیرها) و محور قائم آن شامل فراوانی مطلق (یا فراوانی نسبی و گاهی درصد فراوانی نسبی) است.

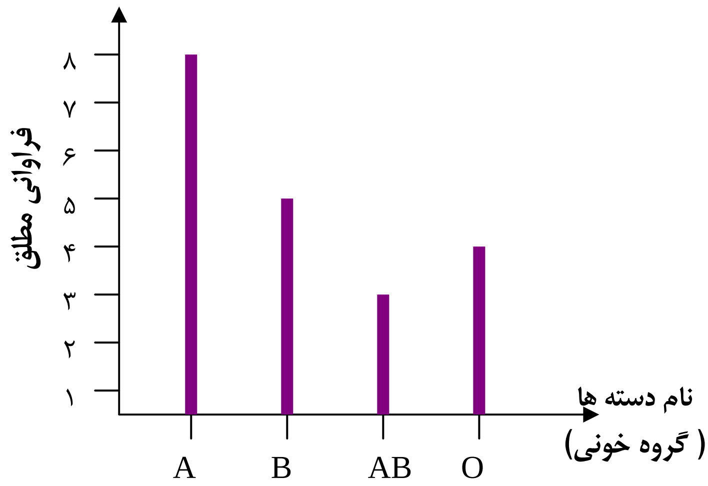


مثال: اطلاعات آماری زیر مربوط به گروه خونی ۲۰ دانش آموز است، پس از تنظیم جدول فراوانی نمودار میله ای رسم کنید.

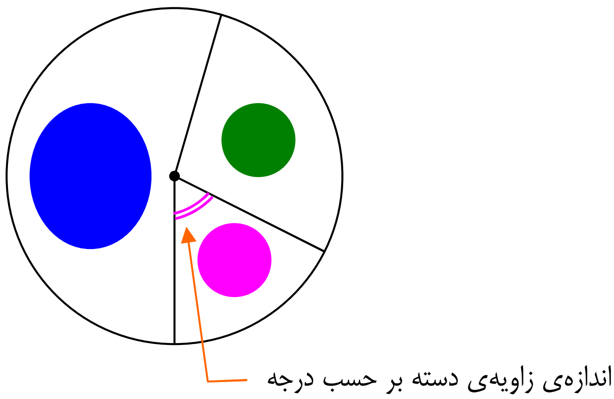
AB,A,AB,A,A,A,B,A,A,B,O,B,O,A,A,B,B,O,AB,O

حل:

دسته ها	A	B	AB	O	جمع
فراوانی	۸	۵	۳	۴	۲۰



نمودار دایره ای: این نمودار نیز برای مقایسه گروه های متفاوت جامعه بکار می‌رود، برای رسم این نمودار زاویه هایی مرکزی (زاویه های قطاع ها) روی دایره طوری مشخص می‌شوند که اندازه‌ی هر زاویه بر حسب درجه به کمک جدول فراوانی به دست آمده باشد.

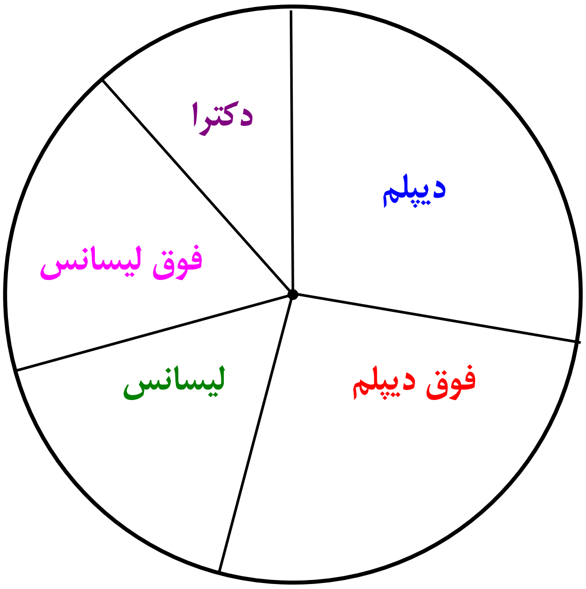


مثال: در جدول زیر مدرک تحصیلی کارکنان یک اداره نوشته شده است. متناظر با این جدول یک نمودار دایره ای رسم کنید.

مدرک تحصیلی	دیپلم	فوق دیپلم	لیسانس	فوق لیسانس	دکترا
فراوانی(تعداد)	۱۰	۹	۶	۷	۴

حل : ابتدا زاویه‌ی قطاع مربوط به هر دسته را تعیین و سپس روی دایره، زاویه های قطاع متناظر را تعیین می کنیم.

نوع بیمه	فراوانی	زاویه‌ی مرکزی (برحسب درجه)
دیپلم	۱۰	۱۰۰
فوق دیپلم	۹	۹۰
لیسانس	۶	۶۰
فوق لیسانس	۷	۷۰
دکترا	۴	۴۰
جمع	۳۶	۳۶۰



تمرین۸ : در یک شرکت دارویی جدول توزیع کارکنان را با نمودار دایره ای نشان می دهیم. زاویه‌ی قطاع مربوط به کارکنان ارشد، چند درجه است؟

نوع مدرک	دیپلم	کاردانی	کارشناسی	ارشد	دکترا
تعداد	۳۰	۹۰	۱۸۰	۱۲۰	۳۰

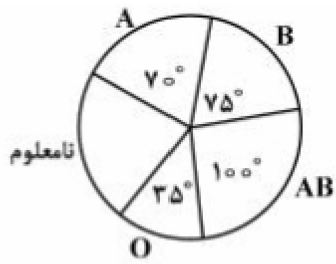
۱) ۸۴° ۲) ۹۲° ۳) ۹۶° ۴) ۱۰۵°

حل :

$$n = ۳۰ + ۹۰ + ۱۸۰ + ۱۲۰ + ۳۰ = ۴۵۰$$

$$\alpha_i = \frac{f_i}{n} \times ۳۶۰ = \frac{۱۲۰}{۴۵۰} \times ۳۶۰ = ۱۲ \times ۸ = ۹۶$$

تمرین۹ : نمودار دایره ای روبرو، متناسب با تعداد کارکنان سازمانی با گروه خونی متمایز است. گروه خونی ۳۲ نفر از آنان تعیین نشده است. چند نفر آنها ، دارای نوع خون B هستند؟



۱) ۲۵ ۲) ۳۰ ۳) ۳۶ ۴) ۴۰

حل : ابتدا زاویه‌ی مربوط به دسته‌ی « نامعلوم» را مشخص می کنیم.

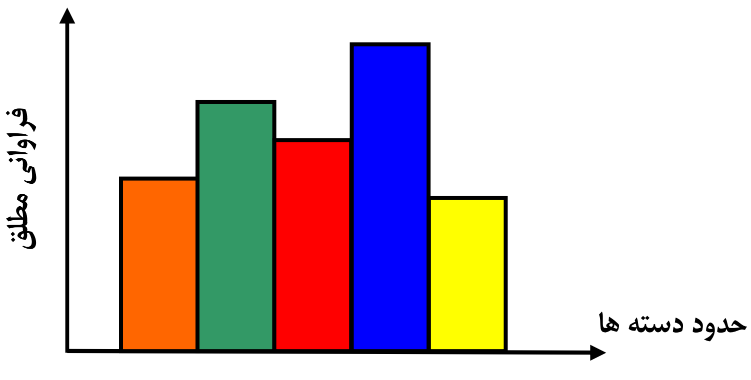
$$۷۰ + ۷۵ + ۱۰۰ + ۳۵ + \alpha = ۳۶۰ \rightarrow \alpha = ۸۰$$

حال از تناسب استفاده می کنیم. بدین شکل که برای دسته‌ی نامعلوم ۸۰ درجه برای ۳۲ نفر است. لذا برای دسته‌ی B زاویه‌ی ۷۵ درجه برای چند نفر است؟

$$\frac{۳۲}{۸۰} = \frac{x}{۷۵} \rightarrow x = \frac{۳۲ \times ۷۵}{۸۰} = ۳۰$$

***

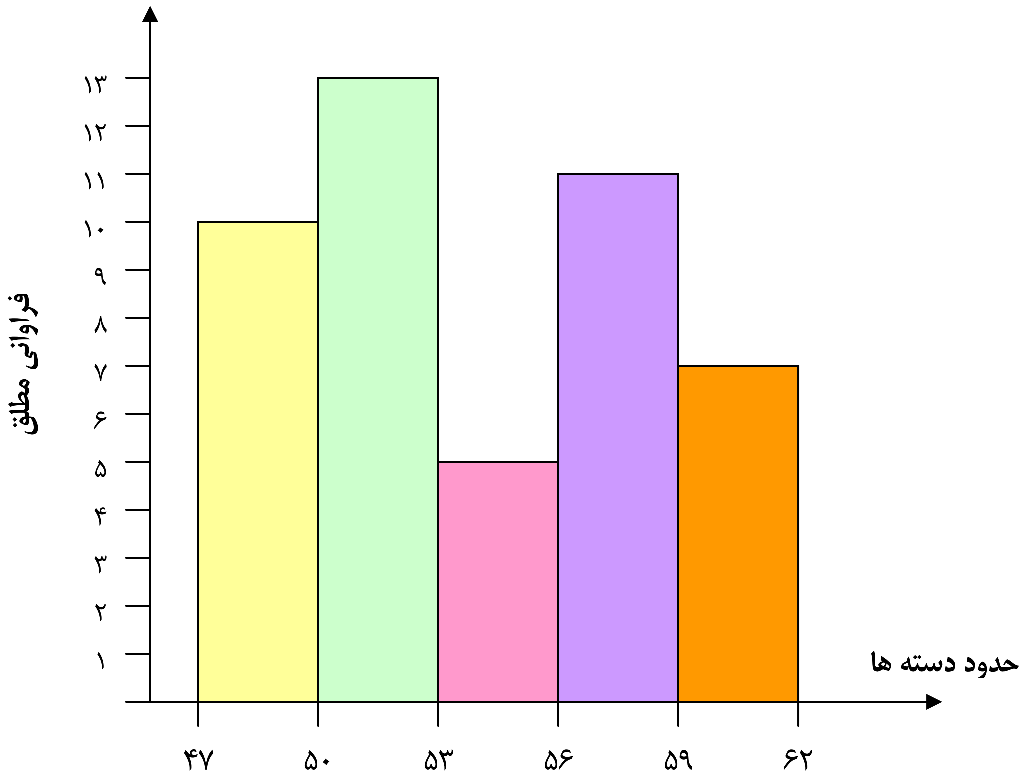
نمودار بافت نگاشت (مستطیلی): در این نمودار مستطیل هایی رسم می شود که قاعده‌ی آنها روی محور افقی و برابر طول هر یک از دسته ها می باشد. بطوری که ارتفاع مستطیل ها به موازات محور قائم و متناسب با فراوانی مطلق یا درصد فراوانی نسبی باشد.



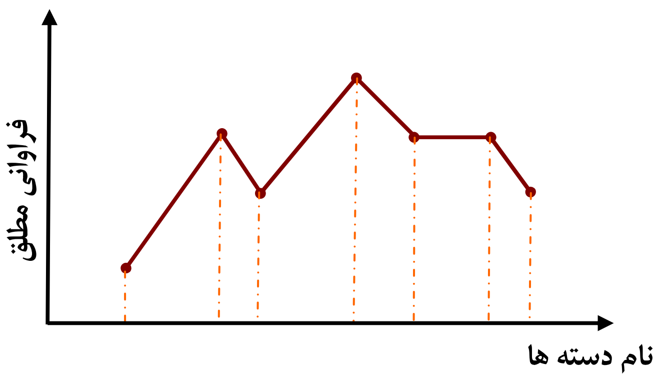
مثال : با توجه به جدول فراوانی زیر یک نمودار مستطیلی رسم کنید.

حدود دسته ها	۴۷–۵۰	۵۰–۵۳	۵۳–۵۶	۵۶–۵۹	۵۹–۶۲
فراوانی	۱۰	۱۳	۵	۱۱	۷

حل:



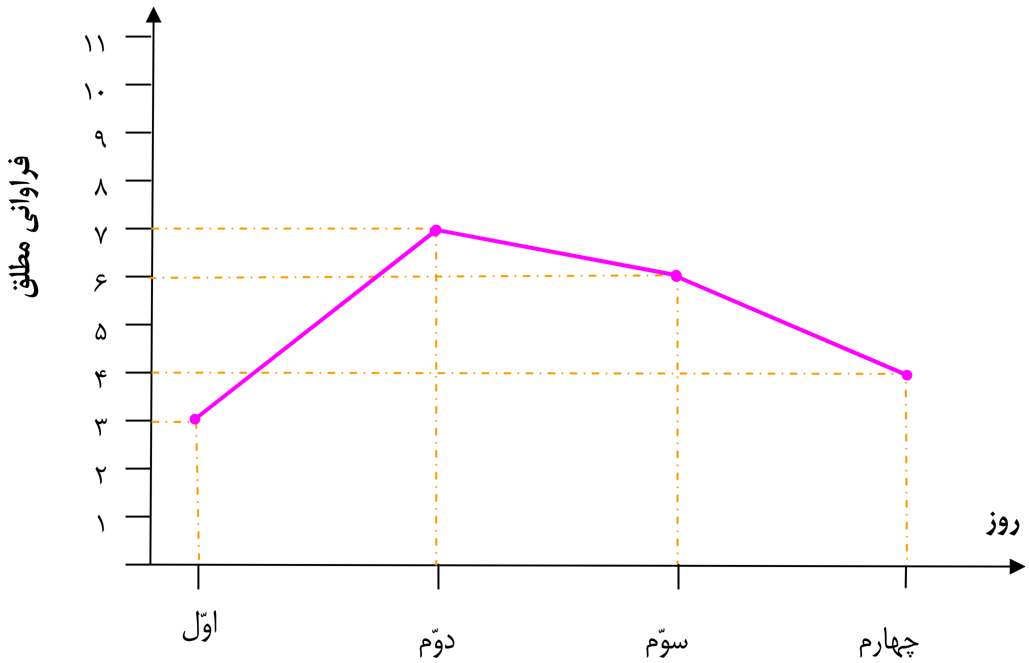
نمودار چندضلعی(چندبر): نمودار چندضلعی جهت نمایش تغییرات داده‌هایی بکار می‌رود که دارای متغیر کمی پیوسته (اغلب زمان) می‌باشند. برای رسم آن روی محور افقی مرکزدسته ها و روی محور قائم فراوانی مطلق یا در صد فراوانی نسبی منظور می شوند.



مثال: جدول زیر تعداد مقالات ارائه شده در یک کنفرانس ریاضی در طول روزهای برگزاری آن است. با توجه به این جدول، نمودار چند ضلعی رسم کنید.

دسته ها	روز اوّل	روز دوّم	روز سوّم	روز چهارم
فراوانی مطلق	۳	۷	۶	۴

حل:



تمرین ۱۰ : با توجه به مثال قبل به سئوالات زیر پاسخ دهید.

الف : کمترین و بیشترین مقالات در چه روز هایی ارائه شده اند؟

ب : چند درصد مقالات در روز سوّم ارائه شده اند؟

تمرین برای حل :

۱۱ : رنگ چشم ۱۲۸ فرد به شرح زیر است.

۶۴ نفر قهوه ای ، ۲۳ نفر آبی ، ۳۶ نفر سبز ، ۵ نفر سایر رنگ ها

الف : کدام نمودار برای این داده ها مناسب است. (میله ای ، دایره ای ، هر دو)

ب : تعیین کنید که چند درصد این افراد رنگ چشم آبی دارند؟

۱۲ : جملات زیر را کامل کنید.

الف : برای متغیرهای پیوسته از نمودار های .... و .... استفاده می شود.

ب : برای متغیرهای گسسته از نمودارهای .... و .... استفاده می شود.

پ : برای متغیرهای کیفی از نمودارهای .... و .... استفاده می شود.

***

درس دوّم : معیارهای گرایش به مرکز

بدیهی است که یکی از اهداف یک فعالیت آماری، نتیجه گیری پیرامون نمونه یا جامعه است. بعد از دسته بندی داده ها و رسم نمودار های مناسب، گاهی محاسبه‌ی معیارهای گرایش به مرکز داده ضروری است. در این درس با مهمترین این معیارها آشنا می شویم. امّا قبل از ورود به بحث نماد جمع و ویژگی های آن را معرفی می کنیم.

معرفی نماد جمع و ویژگی های آن

اگر $x_۱$ و $x_۲$ و $x_۳$ و... و x_n مجموعه ای از n عدد باشند، در این صورت عبارت

$$x_۱ + x_۲ + x_۳ + ... + x_n$$

را به شکل $\sum_{i=۱}^{n} x_i$ نمایش می دهند. به عبارتی دیگر

$$\sum_{i=۱}^{n} x_i = x_۱ + x_۲ + x_۳ + ... + x_n$$

مثال : اگر $x_۱ = ۲$ و $x_۲ = ۳$ و $x_۳ = ۵$ باشد، مقدار $\sum_{i=۱}^{۳} x_i$ را تعیین کنید.

حل :

$$\sum_{i=۱}^{۳} x_i = x_۱ + x_۲ + x_۳ = ۲ + ۳ + ۵ = ۱۰$$

مثال : اگر $x_۱ = ۲$ و $x_۲ = ۳$ و $x_۳ = ۵$ باشد، مقدار $\sum_{i=۱}^{۳} ۴x_i$ را تعیین کنید.

حل :

$$\sum_{i=۱}^{۳} ۴x_i = ۴x_۱ + ۴x_۲ + ۴x_۳ = ۴(x_۱ + x_۲ + x_۳) = ۴(۲ + ۳ + ۵) = ۴۰$$

مثال : اگر $x_۱ = ۲$ و $x_۲ = ۳$ و $x_۳ = ۵$ و $y_۱ = ۷$ و $y_۲ = ۱$ و $y_۳ = ۰$ باشد، مقدار $\sum_{i=۱}^{۳} (x_i + y_i)$ را تعیین کنید.

حل :

$$\sum_{i=۱}^{۳} (x_i + y_i) = (x_۱ + y_۱) + (x_۲ + y_۲) + (x_۳ + y_۳) = (۲ + ۷) + (۳ + ۱) + (۵ + ۰) = ۱۸$$

خواص سیگما

اگر $(x_1$ و x_2 و x_3 و... و $x_n)$ و $(y_1$ و y_2 و y_3 و... و $y_n)$ دو مجموعه‌ی عددی و k یک عدد حقیقی باشد. در این صورت خواص زیر را می توان برای سیگما ثابت نمود.

الف) $\sum_{i=1}^{n} kx_i = k\sum_{i=1}^{n} x_i$

ب) $\sum_{i=1}^{n}(x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

ج) $\sum_{i=1}^{n} k = nk$

تمرین ۱ : هر یک از ویژگی های فوق را ثابت کنید.

***

معیارهای مرکزی

هر عدد که معرف مرکز مجموعه‌ی داده ها باشد را معیار مرکزی یا شاخص مرکزی (پارامتر مرکزی) می نامند. به کمک معیار های مرکزی می توان موقعیت کلی داده ها را تعیین کرد، لذا برای مقایسه‌ی دو یا چند جامعه یک روش مناسب، محاسبه‌ی معیار های مرکزی آنها است. معیار های مرکزی دارای سه نوع مهم هستند. این سه نوع عبارتند از ، میانگین $(\bar{x})$ ، میانه $(\tilde{x})$ و مد یا نما $(\hat{x})$ می باشند که میانگین مهمترین معیار مرکزی محسوب می شود.

***

میانگین

در یک مجموعه‌ی داده های آماری، عدد متمایل به وسط آنها را میانگین (میانگین حسابی) می نامند. میانگین یا متوسط داده ها، با خارج قسمت مجموع اندازه‌ی داده ها بر تعداد آنها برابر است.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

مثال: میانگین داده های زیر را بدست آورید.

x_i : ۵و ۸و۱۰و۱۲و۱۴و۱۷

حل:

$$\bar{x}=\frac{\sum x_i}{n}=\frac{۵+۸+۱۰+۱۲+۱۴+۱۷}{۶}=\frac{۶۶}{۶}=۱۱$$

تمرین۲ : میانگین نمرات یک کلاس۱۶ نفری در درس آمار و احتمال برابر ۱۲/۵ است. جمع نمرات دانش آموزان این کلاس را بدست آورید.

تمرین۳ : میانگین داده های زیر برابر ۲۲ است. مقدار a را بدست آورید.

۲۰ و ۲۵ و۳۰ و ۲۲ و a و ۲۵ و ۲۴ و ۲۱ و ۲۰ و ۱۵

***

میانگین وزنی

گاهی به دلایلی برای داده های یک مجموعه‌ی آماری ضریب های (تعداد تکرار) متفاوتی در نظر گرفته می شود. مانند:

۱ : ضریب های نمرات امتحانات مستمر و پایانی در محاسبه‌ی نمره‌ی سالانه

۲ : ضریب های دروس در کنکور

۳ : واحد های دروس

در این صورت تعداد تکرار هر داده را وزن آن داده می نامند.اگر داده‌های(x_1 و x_2 و x_3 و... و x_n) به ترتیب دارای تعداد تکرار یا ضریب (w_1 و w_2 و w_3 و... و w_n) باشند، می توان میانگین را به روش زیر محاسبه نمود.

داده ها (x_i)	x_1	x_2	x_3	....	x_n
ضریب (w_i)	w_1	w_2	w_3	....	w_n

$$\bar{x}=\frac{w_1x_1+w_2x_2+....+w_nx_n}{w_1+w_2+....+w_n}=\frac{\sum w_ix_i}{\sum w_i}$$

مثال : میانگین داده های مربوط به جدول زیر را محاسبه کنید.

داده ها (x_i)	۱۴	۱۷	۱۶	۱۵
ضریب (w_i)	۴	۲	۳	۱

حل:

$$\bar{x}=\frac{\sum w_i x_i}{\sum w_i}=\frac{۴(۱۴)+۲(۱۷)+۳(۱۶)+۱(۱۵)}{۴+۲+۳+۱}=\frac{۵۶+۳۴+۴۸+۱۵}{۱۰}=\frac{۱۵۳}{۱۰}=۱۵/۳$$

مثال : میانگین داده های زیر را بدست آورید.

۶۲ و۶۵ و۶۵ و۶۵ و۶۵ و۶۲ و۶۲ و۶۰ و۶۲ و۵۸ و ۶۰ و۶۲ و۶۰ و ۵۸ و۷۲

حل:

x_i	F_i	$F_i.x_i$
۷۲	۱	۷۲
۵۸	۲	۱۱۶
۶۰	۳	۱۸۰
۶۲	۵	۳۱۰
۶۵	۴	۲۶۰
جمع	۱۵	۹۳۸

$$\bar{x}=\frac{\sum F_i x_i}{n}=\frac{۹۳۸}{۱۵}\approx ۶۲/۵۳$$

نتیجه : در یک جدول فراوانی اگر x_i مرکز دسته باشد. می توان میانگین داده ها را به کمک فرمول زیر محاسبه کرد.

مثال: با توجه به جدول زیر میانگین حسابی را تعیین کنید.

حدود دسته	۵ – ۱۵	۱۵ – ۲۵	۲۵ – ۳۵	۳۵ – ۴۵
فراوانی	۳	۴	۱	۲

حل :

دسته ها	F_i	x_i	$F_i.x_i$
۵ –۱۵	۳	۱۰	۳۰
۱۵ –۲۵	۴	۲۰	۸۰
۲۵ –۳۵	۱	۳۰	۳۰
۳۵ –۴۵	۲	۴۰	۸۰
جمع	۱۰	-	۲۲۰

$$\bar{x}=\frac{\sum F_i x_i}{n}=\frac{۲۲۰}{۱۰}=۲۲$$

مثال : اگر میانگین داده ها در جدول زیر برابر ۴ باشد، فراوانی مطلق دسته‌ی آخر را به دست آورید.

۶ – ۸	۴ – ۶	۲ – ۴	۰ – ۲	حدود دسته
x	۴	۷	۵	فراوانی

حل :

۷	۵	۳	۱	مرکز دسته
x	۴	۷	۵	فراوانی

$$\bar{x}=\frac{\sum F_i x_i}{\sum F_i}\to ۴=\frac{۵+۲۱+۲۰+۷x}{۵+۷+۴+x}\to ۴(۱۶+x)=۴۶+۷x\to ۳x=۱۸\to x=۶$$

توجه : اگر دو یا چند گروه از یک مجموعه‌ی داده های آماری با اندازه های متفاوت دارای میانگین مشخصی باشند. میانگین کل که به میانگین مرکب معروف است، را می‌توان به روش زیر محاسبه نمود.

$$\bar{z}=\frac{m\bar{x}+n\bar{y}}{m+n}$$

که در آن اندازه و میانگین هر گروه مطابق جدول زیر است.

میانگین	اندازه	–
$\bar{x}$	m	گروه اوّل
$\bar{y}$	n	گروه دوّم

البته این فرمول فوق برای بیش از دو گروه از یک جامعه قابل تعمیم است.

مثال : میانگین ۴ داده‌ی آماری برابر ۷ و میانگین۶ داده‌ی دیگر ۵ است. میانگین کل داده ها را بدست آورید.

حل :

$$\bar{z}=\frac{m\bar{x}+n\bar{y}}{m+n}=\frac{۴(۷)+۶(۵)}{۴+۶}=\frac{۲۸+۳۰}{۱۰}=\frac{۵۸}{۱۰}=۵/۸$$

تمرین۴ : اگر دو گروه دارای اندازه های برابر باشند، میانگین مرکب را از چه روشی می توان محاسبه نمود؟

***

تمرین برای حل :

۵ : در جدول زیر نمره های یک دانش آموز در درس حسابان داده شده است، میانگین وزنی(نمره‌ی سالانه) وی را محاسبه کنید.

پایانی دوّم	مستمر دوّم	پایانی اوّل	مستمر اوّل	–
۱۴	۱۲	۹	۱۰	نمره
۶	۱	۲	۱	ضریب

۶ : دانش آموزی درکنکور سراسری شرکت می کند و نتیجه‌ی کارنامه‌ی آزمون آن به شرح زیر است.

مواد امتحانی	ریاضیات	فیزیک	شیمی	زبان انگلیسی	ادبیات و زبان فارسی	دین و زندگی
درصد	۷۴	۶۵	۸۲	۵۵	۹۵	۱۰۰
ضریب دروس	۴	۳	۲	۱	۲	۳

متوسط درصد مواد امتحانی این دانش آموز را محاسبه کنید.

۷ : سازمانی دارای سه بخش کاری است، تعداد پرسنل و متوسط دریافتی هر بخش بصورت زیر است.

میانگین دریافتی کل پرسنل را محاسبه کنید.

جمع	فرهنگی	آموزشی	اداری	بخش
۱۰۰	۵۰	۳۰	۲۰	تعداد پرسنل
–	۶۰	۴۰	۵۰	متوسط دریافتی

۸ : دستمزد روزانه‌ی ۶ نفر ازکارگران یک کارگاه هر یک ۷۵ هزار تومان و دستمزد روزانه‌ی ۲ نفر دیگر هریک۶۰ هزار تومان است. میانگین دستمزد کل کارگران را حساب کنید.

۹ : میانگین داده های جدول مقابل ۱۴ است، مقدار a را بیابید.

۱۶	۱۲	a	۱۰	داده
۶	۱	۲	۱	فراوانی

۱۰ : دانش آموزی در کنکور سراسری شرکت می کند و نتیجه‌ی کارنامه‌ی آزمون آن به شرح زیر است.

مواد امتحانی	ریاضیات	فیزیک	شیمی	زبان انگلیسی	ادبیات و زبان فارسی	دین و زندگی
درصد	۵۳	؟	۶۷	۳۴	۸۰	۶۷
ضریب دروس	۴	۳	۱	۱	۴	۳

اگر معدل موزون درصد این دانش آموز ۶۳ باشد، درصد این دانش آموز در درس فیزیک را تعیین کنید.

۱۱ : میانگین ۵ داده‌ی آماری ۱۷ است. اگر دو عدد ۱۶ و ۱۱ را به داده های قبلی اضافه کنیم. میانگین جدید چه عددی خواهد شد؟

۱۲ : ثابت کنید که مجموع تفاضل هر یک از داده های یک مجموعه آماری از میانگین آنها برابر صفر است.

$$\sum_{i=۱}^{n}(x_i - \overline{x}) = ۰$$

***

خواص میانگین

اگر $(x_۱ , x_۲ , x_۳ , ... , x_n)$ و $(y_۱ , y_۲ , y_۳ , ... , y_n)$ دو مجموعه‌ی عددی و k یک عدد حقیقی باشد. در این صورت خواص زیر را می توان برای میانگین بررسی کرد.

الف : میانگین حاصل جمع داده های یک مجموعه‌ی آماری با یک عدد ثابت با حاصل جمع آن عدد و میانگین آن داده ها برابر است.

$$z_i = x_i + k \rightarrow \overline{z} = \overline{x} + k$$

ب : میانگین حاصل ضرب داده های یک مجموعه‌ی آماری در یک عدد ثابت با حاصل ضرب آن عدد و میانگین آن داده ها برابر است.

$$z_i = k.x_i \rightarrow \overline{z} = k.\overline{x}$$

ج : میانگین حاصل جمع داده های متناظر از دو مجموعه‌ی داده های آماری با حاصل جمع میانگین های آنها برابر است.

$$z_i = x_i + y_i \rightarrow \overline{z} = \overline{x} + \overline{y}$$

تمرین برای حل :

۱۳ : هر یک از خواص فوق را اثبات کنید.

۱۴ : ثابت کنید که میانگین داده های مساوی، برابر هر یک از آنها است.

۱۵ : الف: اگر $(x_۱ , x_۲ , x_۳ , ... , x_n)$ یک مجموعه از داده های آماری باشند و $y_i = ax_i + b$. ثابت کنید که $\overline{y} = a\overline{x} + b$

ب : هرگاه میانگین داده های آماری $(x_۱ , x_۲ , x_۳ , ... , x_n)$ برابر ۱۷ باشد. میانگین داده های $(۲x_۱ + ۱ , ۲x_۲ + ۱ , ۲x_۳ + ۱ , ... , ۲x_n + ۱)$ را بدست آورید.

تذکر : یکی از کاربرد های خواص میانگین، در محاسبه‌ی سریع میانگین مورد استفاده قرار می گیرد. در این روش یک داده به دلخواه (بهتر است داده‌ی متوسط) را انتخاب کرده و از تمام داده ها کم می کنیم. سپس میانگین داده های جدید را محاسبه می کنیم. با اضافه کردن داده ای که قبلاً از تمام داده ها کم کرده بودیم به میانگین بدست آمده، میانگین واقعی داده ها بدست می آید. به مثال زیر توجه کنید.

مثال: نمرات ریاضی۴۰ دانش آموز یک کلاس در جدول زیر آمده است. میانگین وزنی نمرات، کدام است؟ (کنکور۹۸ ریاضی)

x	۱۰	۱۲	۱۴	۱۵	۱۷	۱۸
f	۵	۸	۷	۱۰	۶	۴

۱) ۱۴/۲ ۲)۱۴/۲۵ ۳)۱۴/۴ ۴) ۱۴/۷۵

حل : از تمام داده ها ۱۴ واحد کم می کنیم و میانگین داده های جدید را به دست می آوریم و در آخر به میانگین به دست آمده ۱۴ واحد را اضافه می کنیم.

$y=x-۱۴$	-۴	-۲	۰	۱	۳	۴
f	۵	۸	۷	۱۰	۶	۴

y	-۴	-۲	۰	۱	۳	۴	جمع
f	۵	۸	۷	۱۰	۶	۴	۴۰
$f\times y$	-۲۰	-۱۶	۰	۱۰	۱۸	۱۶	۸

$$\bar{y}=\frac{\sum(f_i\times y_i)}{\sum f_i}=\frac{۸}{۴۰}=۰/۲$$

$$\bar{y}=\bar{x}-۱۴\rightarrow\bar{x}=\bar{y}+۱۴=۰/۲+۱۴=۱۴/۲$$

تمرین برای حل :

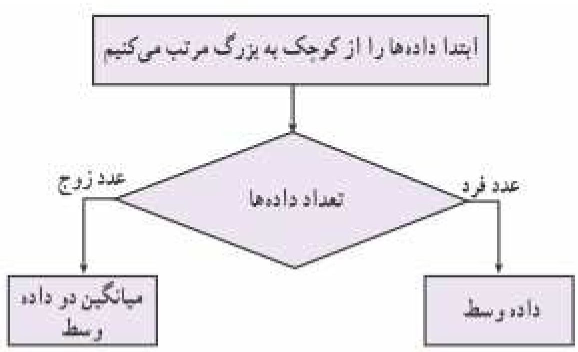
۱۶ : به روش فوق میانگین داده های متناظر با جدول زیر را محاسبه کنید.

x	۱۳	۱۷	۱۶	۱۵
f	۵	۳	۴	۸

***

میانه

در یک مجموعه از داده های آماری، میانه داده ای است که نصف داده از آن بیشتر و نصف داده ها از آن کمتر است. به عبارت دیگر در یک مجموعه‌ی داده های آماری که به صورت غیر نزولی (از کوچک به بزرگ) مرتب شده باشند، عدد وسط این داده ها را میانه می نامند. لذا برای محاسبه‌ی میانه ابتدا داده ها را به صورت غیر نزولی مرتب می‌کنیم، آنگاه



الف) اگر تعداد داده ها فرد باشد، داده‌ی وسط میانه است.

ب) اگر تعداد داده ها زوج باشد، میانگین دو داده‌ی وسط میانه است.

مثال: میانه‌ی هریک از مجموعه های زیر را بیابید.

الف) ۳و ۵ و ۷و ۸و ۹ و۱۰و ۵ و ۸ و ۵

ب) ۱۳ و ۱۵ و۱۶ و۱۷ و۱۷ و ۱۹ و ۲۰ و۲۱ و ۲۲ و۶۰

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

الف:

۳و ۵ و ۵ و ۵ و ۷و ۸ و ۸ و ۹ و ۱۰

چون تعداد داده ها فرد است، پس داده‌ی وسط میانه است. لذا میانه برابر $\tilde{x} = ۷$ می باشد.

ب :

۱۳و۱۵و۱۶و۱۷و۱۷و ۱۹ و ۲۰و۲۱ و۲۲و ۶۰

چون تعداد داده ها زوج است، میانگین دو داده‌ی وسط میانه است است. لذا میانه برابر $\tilde{x} = \frac{۱۷+۱۹}{۲} = ۱۸$ می باشد.

نتیجه: اگر تمام داده های یک مجموعه‌ی داده های آماری برابر باشند، میانه نیز برابر هر یک از آنها است.

توجه: اگر در یک مجموعه‌ی داده های آماری داده‌ی دور افتاده‌ای وجود داشته باشند. چون میانه تحت تأثیر داده های دور افتاده قرار نمی گیرد، نسبت به میانگین، معیار مناسبتری محسوب می شود. داده‌ی دور افتـاده، داده‌ای است که نسبت به سایر داده ها تفاوت بسیار دارد.

***

چارک ها

در یک مجموعه‌ی داده های آماری که داده های آن به صورت غیر نزولی مرتب شده باشند، عدد وسط ایـن داده ها را میانه یا چارک دوّم می نامند و آنرا با $Q_۲$ نمایش می دهند. از طرفی میانه‌ی نیمه‌ی اوّل داده ها را چارک اوّل ($Q_۱$) و میانه‌ی نیمه‌ی دوّم آنها را چارک سوّم ($Q_۳$) می نامند.

مثال: چارک های اوّل تا سوّم داده های زیر را تعیین کنید.

۱۹ و ۳۱ و ۲۵ و ۱۸ و ۳۲ و ۴۳ و ۴۱ و ۳۴ و ۱۸ و ۲۷ و ۱۴ و ۲۳ و ۱۵ و۱۰ و۱۲

حل:

۴۳ و۴۱ و ۳۴ و۳۲ و۳۱ و۲۷ و ۲۵ و ۲۳ و ۱۹ و۱۸ و۱۸ و۱۵ و۱۴ و ۱۲ و۱۰

$Q_۱$ $\qquad$ $Q_۲$ $\qquad$ $Q_۳$

لذا چارک اوّل $Q_۱ = ۱۵$ و چارک دوّم (میانه) $Q_۲ = ۲۳$ و چارک سوّم $Q_۳ = ۳۲$

توجه : برای تعیین چارک ها، ابتدا میانه را محاسبه می کنیم ولی برای تعیین چارک های اوّل و سوّم، میانـه شرکت داده نمی شود.

تمرین۱۷ : چارک های اوّل تا سوّم داده های زیر را محاسبه کنید.

۴۱ و ۱۷ و ۱۶ و ۱۹ و ۱۳ و ۱۴و ۱۸ و ۱۶ و ۱۵ و ۱۷

***

مُد (نما)

در یک مجموعه‌ی داده های آماری هر داده که نسبت به بقیه‌ی داده ها تکرار بیشتری داشته باشد را مُد می نامند. بر همین اساس اگر در یک مجموعه‌ی داده های آماری داده یا داده هایی بیشترین تکرار را داشته باشند آن داده یا داده ها را مد می نامند و در غیر این صورت مجموعه‌ی داده‌ها دارای مد نیست.

مثال: مد هریک از مجموعه های زیر را بیابید.

الف) ۲ و ۵و ۸ و۱۰و۱۵ و ۵و ۲و ۲

ب) ۲ و ۵ و ۸ و ۱۰و ۱۵ و ۲ و ۷ و ۱۵

ج) ۲ و ۵ و ۸ و ۵ و ۲و ۸ و ۸ و ۵ و ۲

د) B و A و B و O و B و B و A و B و AB

حل: الف) $\hat{x}=2$ ب) $\hat{x}=15$ و $\hat{x}=2$ ج) $\hat{x}=$ ندارد د) $\hat{x}=B$

نتیجه:

۱ : با توجه به تعریف هر یک از معیارهای مرکزی، بدیهی است که برای داده‌های کیفی فقط مد را می توان تعیین نمود.

۲ : اگر تمام داده های یک مجموعه‌ی داده های آماری برابر باشند، این مجموعه مد ندارد.

۳ : ممکن است یک مجموعه‌ی داده های آماری دارای بیش از یک مد باشد.

تمرین برای حل :

۱۸ : تعداد حمله های یک تیم فوتبال در شش ماه گذشته به صورت ۴۲ و ۴۴ و ۴۷ و ۱۰ و ۴۳ و ۴۸ است. میانگین و میانه‌ی تعداد حملات این تیم را در این شش ماه به دست آورید. به نظر شما کدام معیار بـا معنـاتر است؟ چرا؟

۱۹: داده های مقابل را در نظر بگیرید. ۲۰ و ۲۴ و ۳۰و ۲۸ و ۲۲و ۲۵ و ۲۲ و ۲۶

الف) میانگین، میانه و مد را محاسبه کنید. ب) چارک های اوّل تا سوّم را تعیین کنید.

۲۰: توضیح دهید، اگرتمام داده های یک مجموعه‌ی آماری با عدد ثابت۳ جمع شوند، میانه و مد آنها چه تغییری می‌کنند؟

۲۱: توضیح دهید اگر تمام داده های یک مجموعه‌ی آماری درعدد ثابت ۲ ضرب شوند، میانه و مد آنها چه تغییری می کنند؟

۲۲: یک کارخانه تولید خودرو برای رنگ اتومبیل از مشتریان نظرخواهی کرد و جدول زیر را از یک نمونه بدست آورد. با توجه به این جدول مد داده ها را تعیین کنید.

زرد	آبی	قرمز	سفید	سبز	رنگ مورد علاقه مشتری
۵	۱۷	۱۳	۴۲	۱۵	تعداد (نفر)

۲۳: ثابت کنید که اگر داده های یک مجموعه‌ی داده های آماری تشکیل دنباله‌ی حسابی می دهند، میانگین آنها برابر $\frac{a+b}{۲}$ است که در آن a جمله‌ی اوّل و b جمله‌ی آخر دنباله است.

۲۴ : در هر مورد تعیین کنید که کدام معیار گرایش به مرکز، مناسبتر است.

الف : نتایج رأی گیری در انتخاب ریاست جمهوری

ب : تعیین میزان دمای یک شهر در یک هفته

ج : تعیین طول عمر لامپ های کم مصرف

د : تعیین حد وسط حقوق کارکنان چند اداره با احتساب کلیه‌ی حقوق ها

۲۵ : داده های زیر را در نظر بگیرید.

۶ و ۹ و ۱۱ و ۱۳ و ۵ و ۴ و ۶ و ۱۰

الف : میانگین، میانه و مد را محاسبه کنید.

ب : یک داده از این داده ها را طوری تغییر دهید که مد و میانه ثابت بماند.

ج : دو داده‌ی دیگر به این داده ها، طوری اضافه کنید که میانگین تغییر نکند.

د : دو داده‌ی دیگر به این داده ها، طوری اضافه کنید که میانگین و میانه ثابت بمانند ولی مد تغییرکند.

۲۶ : داده های زیر را طوری تغییر دهید که میانه ثابت بماند ولی میانگین دو واحد اضافه شود.

۲۲ و ۱۷ و ۱۲ و ۱۹ و ۱۰

۲۷ : در هر مورد جای خالی را کامل کنید.

الف : اگر در یک مجموعه‌ی داده های آماری، داده‌ی دورافتاده وجود داشته باشد، .... نسبت به میانگین، معیار بهتری محسوب می شود.

ب : اگر در مجموعه‌ی داده های آماری، تمام داده ها .... باشند. این داده ها مُد ندارند.

***

تهیه کننده : جابر عامری

عضو گروه ریاضی دوره ی دوم متوسطه

استان خوزستان

درس سوّم : معیارهای پراکندگی

در فعالیت های آماری، گاهی لازم می شود که میزان پراکندگی (دوری و نزدیکی) داده را نسبت به همدیگر یا نسبت به میانگین محاسبه شود. برای اینکار از معیارهای پراکنـدگی اسـتفاده مـی شـود. در ایـن درس بـا مهمترین این معیارها آشنا می شویم.

معیار های پراکندگی

هر عدد که میزان پراکندگی داده ها نسبت به همدیگر یا نسبت به میانگین را نشان دهد، را معیار پراکندگی یا شاخص پراکندگی (پارامتر پراکندگی) می نامند. مهمترین معیارهای پراکندگی عبارتند از ، دامنه‌ی تغییـرات ، واریانس ، انحراف معیار و ضریب تغییرات می باشند.

از بین این معیارها، دامنه‌ی تغییرات پراکندگی داده ها را نسبت به همدیگر و بقیه، پراکندگی را نسبت به میانگین که مهمترین معیار مرکزی می‌باشد، نشان می دهند و توجه داشته باشید که انحراف معیار به جهت ویژگی هایی که دارد ، مهمترین معیار محسوب می‌شود.

دامنه‌ی تغییرات

در یک مجموعه‌ی داده های آماری تفاضل کمترین داده از بیشترین آنها را دامنه‌ی تغییـرات مـی نامنـد. بـه عبارتی دیگر دامنه‌ی تغییرات، طول بازه ای اسـت کـه داده هـا در آن قـرار دارنـد. در ایـن صـورت، اگـر a کوچکترین و b بزرگترین داده‌ی یک مجموعه‌ی داده های آماری باشند. دامنه‌ی تغییرات به شکل زیر است.

$$R = b - a$$

دامنه‌ی تغییرات را می توان به بیشترین اختلاف بین داده های یک مجموعه‌ی داده های آماری تعبیر کرد.

مثال: دامنه‌ی تغییرات داده های زیر را حساب کنید و تعبیر آن را بنویسید.

۱۳ و ۸ و ۱۴ و ۱۵ و۱۰

حل:

$$R = b - a = ۱۵ - ۸ = ۷$$

تعبیر: بیشترین اختلاف بین داده های این مجموعه برابر ۷ است.

نتیجه:

۱ : بزرگی دامنه‌ی تغییرات نشان دهنده‌ی تفاوت زیاد در جامعه است، هر چه قدر این دامنه بیشتر باشد، تفاوت بین داده ها زیاد است و هر چه قدر این دامنه کمتر باشد، داده ها به هم نزدیکترند. اگر دامنه‌ی تغییرات صفر باشد، تمام داده ها برابر هستند و جامعه همگون است.

۲ : دامنه‌ی تغییرات ضعیف ترین شاخص پراکندگی است و معرف خوبی برای پراکندگی داده ها نمی باشد، زیرا برای محاسبه‌ی آن فقط از بزرگترین و کوچکترین داده استفاده می شود و تعداد یا مقادیر بقیه‌ی داده‌ها تأثیری بر مقدار آن ندارند.

***

واریانس و انحراف معیار

میانگین توان دوّم تفاضل داده ها از میانگین آنها را واریانس (پراش) می نامند. در این صورت

$$\sigma^{2}=\frac{\sum(x_i-\overline{x})^{2}}{n}$$

توجه داشته باشید که واریانس در ضمن داشتن اهمیت زیاد دارای دو اشکال عمده است.

الف : تحت تأثیر داده های بزرگ قرار می گیرد.

ب : واحد اندازه گیری واریانس مجذور واحد اصلی متغیر داده ها است. مثلاً اگر واحد اندازه گیری داده ها سانتی متر باشد، واحد اندازه گیری واریانس سانتی متر مربع خواهد بود.

برای رفع این دو اشکال از انحراف معیار استفاده می شود. ریشه‌ی دوّم واریانس را انحراف معیار (انحراف استاندارد) می نامند.

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

بنابراین طبق این تعریف بدیهی است که انحراف معیار دارای همان واحدی خواهد بود که داده ها بر حسب آن محاسبه شده اند.

مثال : واریانس و انحراف معیار داده های زیر را بدست آورید.

۹ و۴ و۷ و ۳ و ۲

حل:

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^۲$
۲	-۳	۹
۳	-۲	۴
۷	۲	۴
۴	-۱	۱
۹	۴	۱۶
جمع $= ۲۵$	-	۳۴

میانگین $\bar{x} = \frac{\sum x_i}{n} = \frac{۲۵}{۵} = ۵$

واریانس $\sigma^۲ = \frac{\sum (x_i - \bar{x})^۲}{n} = \frac{۳۴}{۵} = ۶/۸$

انحراف معیار $\sigma = \sqrt{\sigma^۲} = \sqrt{۶/۸} \approx ۲/۶$

توجه: اگر مجموعه‌ی داده های آماری دارای تکرار باشند، می توان به کمک فرمول زیر واریانس را محاسبه کرد.

$$\sigma^۲ = \frac{\sum F_i (x_i - \bar{x})^۲}{n}$$

همچنین اگر x_i مرکز دسته ها در یک جدول فراوانی باشد. می توان به کمک این فرمول واریانس داده ها را نیز به دست آورد.

مثال : واریانس و انحراف معیار داده های زیر را بدست آورید.

۱۲ و ۱۵ و ۱۲ و ۱۲ و ۱۲ و ۱۲ و ۱۱ و ۱۱ و ۱۱ و۱۰ و ۱۰ و ۸ و۷ و۷ و۷ و

حل:

x_i	F_i	$F_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^۲$	$F_i (x_i - \bar{x})^۲$
۷	۳	۲۱	۳−	۹	۲۷
۸	۱	۸	۲−	۴	۴
۱۰	۲	۲۰	۰	۰	۰
۱۱	۳	۳۳	۱	۱	۳
۱۲	۵	۶۰	۲	۴	۲۰
۱۵	۱	۱۵	۵	۲۵	۲۵
جمع	۱۵	۱۵۷	−	−	۷۹

میانگین $\bar{x} = \frac{\sum F_i x_i}{n} = \frac{۱۵۷}{۱۵} = ۱۰/۴۶ \approx ۱۰$

واریانس $\sigma^۲ = \frac{\sum F_i (x_i - \bar{x})^۲}{n} = \frac{۷۹}{۱۵} = ۵/۲۶$

انحراف معیار $\sigma = \sqrt{\sigma^۲} = \sqrt{۵/۲۶} \approx ۲/۲۹$

تمرین برای حل :

۱ : سن ۲۰ نفر شرکت کننده در یک کلاس نهضت سواد آموزی به شرح زیر است.

۱۲ و ۲۱ و۱۹ و ۲۴ و ۲۱ و ۲۳ و ۲۷ و ۱۳ و ۲۳ و ۱۷ و ۲۵ و ۱۲ و ۲۷ و ۲۲ و ۲۲ و ۲۵ و ۱۷ و۱۳ و۱۸ و۲۷

الف : دامنه‌ی تغییرات داده ها را به دست آورید.

ب : جدول زیر را با توجه به این داده ها تکمیل نموده و سپس انحراف معیار آنها را به دست آورید.

دسته ها	F_i	x_i	$F_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^۲$	$F_i (x_i - \bar{x})^۲$
۱۵− ۱۰						
۲۰− ۱۵						
۲۵− ۲۰						
۳۰− ۲۵						
جمع						

۲ : ثابت کنید که واریانس داده های مساوی برابر صفر است.

***

خواص واریانس

اگر ($x_۱$ و $x_۲$ و $x_۳$ و... و x_n) یک مجموعه‌ی عددی و k یک عدد حقیقی باشد. در این صورت خواص زیر را می توان برای واریانس آنها بررسی کرد.

۱ : واریانس حاصل جمع داده ها با یک عدد ثابت ، با واریانس آن داده ها برابراست.

$$y_i = x_i + k \to \sigma_y^۲ = \sigma_x^۲$$

۲ : واریانس حاصل ضرب داده ها در یک عدد ثابت ، با حاصل ضرب مربع آن عدد در واریانس آن داده ها برابر است.

$$y_i = k.x_i \to \sigma_y^۲ = k^۲.\sigma_x^۲$$

تمرین۲ : هر یک از خواص فوق را اثبات کنید.

***

خواص انحراف معیار

مشابه آنچه که برای واریانس داشتیم. خواص زیر را برای انحراف معیار نیز می توان بیان کرد.

۱ : انحراف معیار حاصل جمع داده ها با یک عدد ثابت با انحراف معیارآن داده ها برابراست.

$$y_i = x_i + k \to \sigma_y = \sigma_x$$

۲ : انحراف معیار حاصل ضرب داده ها در یک عدد ثابت، با حاصل ضرب قدر مطلق آن عدد در انحراف معیارآن داده ها برابر است.

$$y_i = k.x_i \to \sigma_y = |k|.\sigma_x$$

تمرین۳ : هر یک از خواص فوق را اثبات کنید.

تمرین۴ : الف: اگر ($x_۱$ و $x_۲$ و $x_۳$ و... و x_n) یک مجموعه از داده های آماری باشند و $y_i = ax_i + b$.

ثابت کنید که $\sigma_y^۲ = a^۲.\sigma_x^۲$

ب : هرگاه واریانس داده های آماری($x_۱$ و $x_۲$ و $x_۳$ و... و x_n) برابر۱۶ باشد. انحراف معیار داده های ($۲x_۱+۱$ و $۲x_۲+۱$ و $۲x_۳+۱$ و... و $۲x_n+۱$) را محاسبه کنید.

تمرین۵ : اگر واریانس داده های یک مجموعه‌ی داده های آماری برابر ۱۸ و میانگین آنها ۸ باشد و تمام داده ها را دو برابر کنیم. میانگین و واریانس داده های جدید را تعیین کنید.

***

روش دیگر برای محاسبه‌ی واریانس

با توجه به تعریف واریانس و خواص سیگما می توان نتیجه گرفت که :

$$\sigma^۲ = \frac{\sum x_i{}^۲}{n} - \bar{x}^۲$$

برای اثبات این موضوع می توان به شکل زیر عمل کرد.

$$\sigma^۲ = \frac{\sum (x_i - \bar{x})^۲}{n}$$

$$\rightarrow \sigma^۲ = \frac{\sum (x_i{}^۲ - ۲\bar{x}x_i + \bar{x}^۲)}{n} = \frac{\sum x_i{}^۲ + \sum(-۲\bar{x}x_i) + \sum \bar{x}^۲}{n}$$

$$\rightarrow \sigma^۲ = \frac{\sum x_i{}^۲}{n} + \frac{\sum(-۲\bar{x}x_i)}{n} + \frac{\sum \bar{x}^۲}{n} = \frac{\sum x_i{}^۲}{n} - ۲\bar{x}\frac{\sum x_i}{n} + \frac{n\bar{x}^۲}{n}$$

$$\rightarrow \sigma^۲ = \frac{\sum x_i{}^۲}{n} - ۲\bar{x}.\bar{x} + \bar{x}^۲ = \frac{\sum x_i{}^۲}{n} - ۲\bar{x}^۲ + \bar{x}^۲ = \frac{\sum x_i{}^۲}{n} - \bar{x}^۲$$

نتیجه : اگر داده های آماری دارای تکرار باشند. می توان به شکل زیر نیز عمل کرد.

$$\sigma^۲ = \frac{\sum F_i x_i{}^۲}{n} - \bar{x}^۲$$

مثال : واریانس و انحراف معیار داده های زیر را به روش فوق محاسبه کنید.

۷ و ۹ و۶ و ۵ و۳

حل :

x_i	$x_i^{۲}$
۳	۹
۵	۲۵
۶	۳۶
۹	۸۱
۷	۴۹
جمع =۳۰	۲۰۰

میانگین $\bar{x} = \frac{\sum x_i}{n} = \frac{۳۰}{۵} = ۶$

واریانس $\sigma^{۲} = \frac{\sum x_i^{۲}}{n} - \bar{x}^{۲} = \frac{۲۰۰}{۵} - (۶)^{۲} = ۴۰ - ۳۶ = ۴$

انحراف معیار $\sigma = \sqrt{۴} = ۲$

تمرین برای حل :

۶ : اگر برای یک جامعه‌ی داده های آماری داشته باشیم.

$\sum x_i^{۲} = ۵۸۰$ و $\sum x_i = ۱۰۰$ و $n = ۲۰$

واریانس و انحراف معیار داده ها را تعیین کنید.

***

واریانس مرکب[1]

اگر دو یا چند گروه از یک مجموعه‌ی داده های آماری با اندازه های متفاوت دارای واریانس های متفاوتی باشند، واریانس کل که به واریانس مرکب معروف است را می‌توان از روش زیر محاسبه کرد.

$$\sigma_t^{۲} = \frac{m(\bar{x}^{۲} + \sigma_x^{۲}) + n(\bar{y}^{۲} + \sigma_y^{۲})}{m+n} - \bar{z}^{۲}$$

1. فقط برای مطالعه

با فرض اینکه این دو گروه که دارای میانگین و واریانس مطابق جدول زیر می باشند.

واریانس	میانگین	اندازه	–
$\sigma_x^۲$	$\bar{x}$	m	گروه اوّل
$\sigma_y^۲$	$\bar{y}$	n	گروه دوّم

همچنین $\bar{z}$ میانگین مرکب دو گروه است.

توجه داشته باشید که

الف : با ریشه گیری از فرمول فوق انحراف معیار مرکب نیز بدست می آید.

ب : این فرمول، برای بیش از دو گروه از یک جامعه نیز قابل تعمیم است.

مثال: اطلاعات زیر مربوط به اجرای یک آزمون هوش در دو کلاس است.

انحراف معیار	میانگین	اندازه	–
۹	۱۰۳	۲۵	کلاس اوّل
۱۲	۱۰۴	۲۴	کلاس دوّم

واریانس و انحراف معیار مرکب را محاسبه کنید.

حل :

$$\text{میانگین}\ \bar{z}=\frac{m\bar{x}+n\bar{y}}{m+n}=\frac{۲۵(۱۰۳)+۲۴(۱۰۴)}{۲۵+۲۴}=\frac{۵۰۷۱}{۴۹}\approx ۱۰۳/۵$$

$$\text{واریانس}\ \sigma_t^۲=\frac{m(\bar{x}^۲+\sigma_x^۲)+n(\bar{y}^۲+\sigma_y^۲)}{m+n}-\bar{z}^۲$$

$$=\frac{۲۵(۱۰۳^۲+۹^۲)+۲۴(۱۰۴^۲+۱۲^۲)}{۲۵+۲۴}-(۱۰۳/۵)^۲=\frac{۲۶۷۲۵۰+۲۶۳۰۴۰}{۴۹}-۱۰۷۱۲/۲۵=۱۱۰$$

$$\text{انحراف معیار}\ \sigma_t=\sqrt{\sigma_t^۲}=\sqrt{۱۱۰}\approx ۱۰/۴۸$$

تمرین ۷ : اگر اندازه های دو گروه برابر باشند، واریانس مرکب از چه رابطه ای به دست می آید؟

***

ضریب پراکندگی

خارج قسمت انحراف معیار داده های یک مجموعه‌ی داده های آماری بر میانگین آنها را ضریب پراکندگی (ضریب تغییرات) گویند.

$$CV_x = \frac{\sigma}{\overline{x}}$$

توجه کنید که طبق تعریف ، ضریب پراکندگی داده ها فاقد واحد اندازه گیری می‌باشد، لذا اغلب در موارد زیر استفاده می شود.

الف : برای مقایسه‌ی دو یا چند جامعه‌ی آماری که واحد اندازه گیری داده‌های آنها متفاوت باشد.

مثلاً: مقایسه‌ی سود دو شرکت که سود یکی برحسب ریال و سود دیگری بر حسب دلار باشد.

ب : برای مقایسه‌ی دو یا چند جامعه‌ی آماری که واریانس های آنها برابر باشند ولی میانگین‌های متفاوت دارند.

مثلاً: باتوجه به اطلاعات زیر دو شرکت x و y دارای انحراف معیار برابر هستند ولی میانگین مساوی دارند، لذا شرکت y از پراکندگی شرکت بهتری در مقایسه با شرکت x محسوب می شود.

ضریب پراکندگی	انحراف معیار	میانگین	–
۰/۲	۲	۱۰	شرکت x
۰/۱	۲	۲۰	شرکت y

مثال : ضریب پراکندگی داده های زیر را محاسبه کنید.

۶ و ۸ و ۳ و ۹ و ۴

$(x_i - \overline{x})^۲$	$x_i - \overline{x}$	x_i
۴	−۲	۴
۹	۳	۹
۹	−۳	۳
۴	۲	۸
۰	۰	۶
۲۶	---	جمع = ۳۰

میانگین $\overline{x} = \frac{\sum x_i}{n} = \frac{۳۰}{۵} = ۶$

واریانس $\sigma^{۲}=\frac{\sum(x_i-\overline{x})^{۲}}{n}=\frac{۲۶}{۵}=۵/۲$

انحراف معیار $\sigma=\sqrt{۵/۲}\approx ۲/۲۸$

ضریب تغییرات $CV=\frac{\sigma}{\overline{x}}=\frac{۲/۲۸}{۶}\approx ۰/۳۸$

نتیجه: اگر تمام داده های آماری برابر باشند، ضریب تغییرات آنها صفر است.

مثال : فرض کنیم وزن دو نوزاد ۱/۵ و ۲/۵ کیلوگرم و وزن دو فرد چهل ساله ۸۰ و ۸۱ کیلوگرم است.

الف : تفاوت وزن دو نوزاد چقدر است.

ب : تفاوت وزن دو فرد چهل ساله چقدر است.

پ : انحراف معیارهای وزن هر دو دسته را به دست آورید.

ت : فکر می کنید تفاوت وزن ها در کدام دسته زیادتر به نظر می رسد؟

ث : ضریب تغییرات هر دو دسته را به دست آورید. چه نتیجه ای می گیرید؟

حل :

تفاوت وزن فرد چهل ساله $R_y=۸۱-۸۰=۱$ تفاوت وزن دو نوزاد $R_x=۲/۵-۱/۵=۱$

x_i	$(x_i-\overline{x})^{۲}$
۱/۵	$(۱/۵-۲)^{۲}=۰/۲۵$
۲/۵	$(۲/۵-۲)^{۲}=۰/۲۵$
$\overline{x}=\frac{۱/۵+۲/۵}{۲}=۲$	$\sigma_x=\sqrt{\frac{۰/۲۵+۰/۲۵}{۲}}=۰/۵$

ضریب تغییرات

$C.V_x=\frac{\sigma_x}{\overline{x}}=\frac{۰/۵}{۲}=۰/۲۵$

y_i	$(y_i-\overline{y})^{۲}$
۸۰	$(۸۰-۸۰/۵)^{۲}=۰/۲۵$
۸۱	$(۸۱-۸۰/۵)^{۲}=۰/۲۵$
$\overline{y}=\frac{۸۰+۸۱}{۲}=۸۰/۵$	$\sigma_y=\sqrt{\frac{۰/۲۵+۰/۲۵}{۲}}=۰/۵$

ضریب تغییرات

$C.V_y=\frac{\sigma_y}{\overline{y}}=\frac{۰/۵}{۸۰/۵}=۰/۰۰۶$

چون انحراف معیار هر دو گروه برابر است. نمی توان گفت که تفاوت وزن ها در کدام دسته بیشتر است. لذا بهتر است از ضریب تغییرات استفاده کرد. گروهی که ضریب تغییرات آن کمتر است، تفاوت وزن های کمتر دارد.

تمرین برای حل :

۸ : برای داده های مقابل ضریب تغییرات را محاسبه کنید. ۱۲ و ۱۵ و ۳ و ۶ و ۹

۹ : ضریب تغییرات داده های مربوط به جدول زیر را تعیین کنید.

داده ها	۱	۲	۴	۵
فراوانی	۵	۱	۲	۴

۱۰ : اگر ضریب تغییرات ۱۰ داده برابر ۲ و میانگین آنها برابر ۴ باشد، واریانس داده ها را به دست آورید.

۱۱ : اگر $n=۱۰$ و $\sum x_i = ۶۰$ و $\sum {x_i}^۲ = ۴۰۰$ ضریب پراکندگی داده ها را تعیین کنید.

۱۲ : نشان دهید که اگر تمام داده های یک مجموعه‌ی داده های آماری را در عدد مثبت k ضرب کنیم، ضریب تغییرات آنها تغییری نمی کند.

۱۳ : توضیح دهید که اگر تمام داده های یک مجموعه‌ی داده های آماری را با عدد مثبت k جمع کنیم، ضریب تغییرات آنها چه تغییری می کند؟

۱۴ : داده های یک مجموعه ی داده های آماری بر حسب متر اندازه گیری شده اند. واحد اندازه گیری معیارهای میانگین ، واریانس ، انحراف معیار و ضریب تغییرات را مشخص کنید.

۱۵ : اگر همه‌ی داده های یک مجموعه‌ی داده های آماری مساوی باشند. درباره‌ی میانگین و واریانس آن مجموعه چه می توان گفت؟

۱۶ : اگر تمام داده های یک مجموعه‌ی آماری مساوی و میانگین آنها ۵ باشد. مقدار هر یک از شاخص های زیر را در صورت وجود بنویسید. الف : مد ب : میانه ج : انحراف معیار

۱۷ : میانگین یک مجموعه‌ی داده های آماری ۵ و واریانس آنها ۱۲ است. در صورتی که تمام داده ها را سه برابر کنیم، میانگین و واریانس داده های جدید را بنویسید.

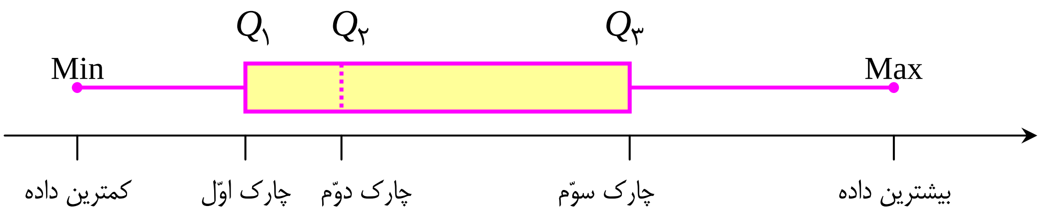
۱۸ : امتیازات مهارت کاری دو نفر از پرسنل یک آزمایشگاه در ۵ روز کاری به صورت زیر است. تعیین کنید که دقت عمل کدام یک بیشتر است.

A : ۲۰ و ۲۳ و ۲۴ و ۲۶ و ۲۷ و B : ۲۱ و ۲۲ و ۲۳ و ۲۶ و ۲۸

***

نمودار جعبه ای و دامنه‌ی میان چارکی

نموداری که میزان پراکندگی داده ها را می توان به کمک آن تعیین کرد. نموداری موسوم به نمودار جعبه ای است.این نمودار پراکندگی داده را در اطراف میانه و قبل و بعد از چارک اوّل و سوّم را به تصویر می کشد. این نمودار به کمک کمترین و بیشترین داده و چارک های اوّل تا سوّم داده های یک مجموعه‌ی آماری ترسیم می شود.



در نمودار جعبه ای ممکن است، میانه وسط جعبه نباشد. میانه به هر سمت متمایل شود، در آن سمت داده ها به هم نزدیکترند و در سمت دیگر داده ها پراکنده تر هستند.

یک معیار دیگر که در صورت وجود داده های دور افتاده (پَرت) مورد استفاده قرار مـی گیـرد. دامنـه‌ی میـان چارکی است. تفاضل چارک اوّل از چارک سوّم را دامنه‌ی میان چارکی می نامند و آن را با IQR نمایش می دهند.

$$IQR = Q_{۳} - Q_{۱}$$

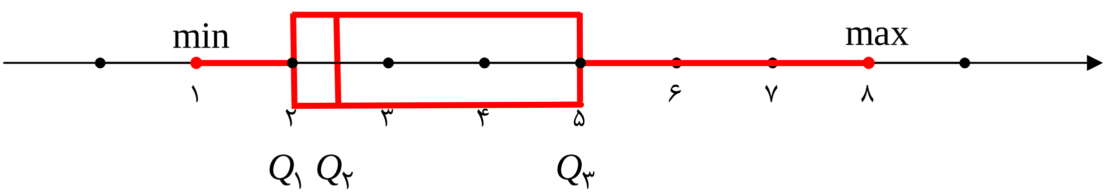
مثال: حسین تعداد گلهای ۲۲ بوته‌ی گل قرمز یک باغ را شمرده است. نمودار جعبه ای این داده ها را ترسیم نموده و دامنه‌ی میان چارکی را محاسبه کنید.

۸ و ۷ و ۷ و ۶ و ۵ و ۵ و ۴ و ۴ و ۴ و ۳ و ۳ و ۲ و ۲ و ۲ و ۲ و ۲ و۲ و ۱ و۱ و ۱ و ۱ و ۱

حل :

$\boxed{۸ و ۷ و ۷ و ۶ و ۵ و ۵ و ۴ و ۴ و ۴ و ۳ و ۳}$ و $\boxed{۲ و ۲ و ۲ و ۲ و ۲ و۲ و ۱ و ۱ و۱ و ۱ و ۱}$

$$\min = ۱ \text{ و } Q_{۱} = ۲ \text{ و } Q_{۲} = \frac{۲+۳}{۲} = ۲/۵ \text{ و } Q_{۳} = ۵ \text{ و } \max = ۸$$



دامنه‌ی میان چارکی $IQR = Q_{۳} - Q_{۱} = ۵ - ۲ = ۳$

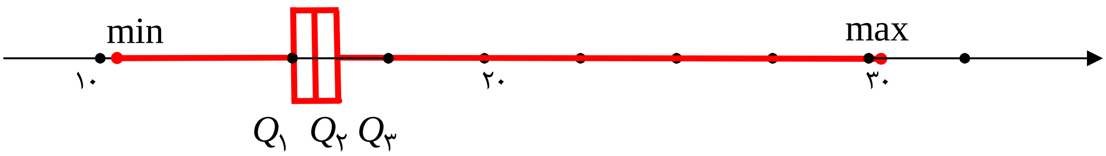
مثال : داده های زیر نرخ بیکاری یک کشور در ده سال گذشته است.

۱۱/۵ و ۱۱/۳ و ۱۰/۵ و ۱۰/۴ و ۱۱/۹ و ۱۳/۵ و ۱۲/۳ و ۱۲/۲ و ۱۰/۲ و ۳۰/۱

نمودار جعبه ای متناظر با این داده ها را رسم کنید.

$\boxed{\text{۳۰/۱ و ۱۳/۵ و ۱۲/۳ و ۱۲/۲ و ۱۱/۹}}$ و $\boxed{\text{۱۱/۵ و ۱۱/۳ و ۱۰/۵ و ۱۰/۴ و ۱۰/۲}}$

$$\min = ۱۰/۲ \text{ و } Q_۱ = ۱۰/۵ \text{ و } Q_۲ = \frac{۱۱/۵ + ۱۱/۹}{۲} = ۱۱/۷ \text{ و } Q_۳ = ۱۲/۳ \text{ و } \max = ۳۰/۱$$



مثال : در بررسی سن بازیکنان تیم ملّی،کدام نمودار مناسبتر است.

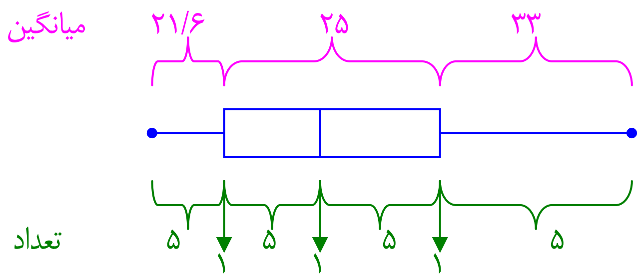
الف) دایره ای　　ب) میله ای　　ج) جعبه ای　　د) بافت نگاشت

حل : با توجه به اینکه نمودار جعبه ای، پراکندگی سن بازیکنان در اطراف میانه را به خوبی نشان می دهد، نمودار جعبه ای مناسبتر است.

مثال : در نمودار جعبه ای ۲۳ داده‌ی آماری ، میانگین دنباله های سمت چپ و سمت راست به ترتیب ۲۱/۶ و ۳۳ و میانگین داده های داخل و روی جعبه ۲۵ می باشد. میانگین کل این داده ها کدام است؟

۱) ۲۵/۸　　۲) ۲۶　　۳) ۲۶/۱　　۴) ۲۶/۲

حل : با کمی دقت می توان برای مسئله نمودار زیر را رسم نمود.



در این صورت میانگین کل به صورت زیر است.

$$\bar{x} = \frac{۵(۲۱/۶) + ۱۳(۲۵) + ۵(۳۳)}{۲۳} = \frac{۵۹۸}{۲۳} = ۲۶$$

تمرین برای حل :

۱۹ : داده های زیر را در نظر بگیرید.

۱۹و ۳۱ و ۲۵ و ۱۸ و ۳۲ و ۴۳ و ۴۱ و ۳۴ و ۱۶ و ۲۷ و ۱۴ و ۲۳ و ۱۵ و ۱۰ و ۱۲

الف : چارک های اوّل تا سوّم داده های زیر را تعیین کنید.

ب : دامنه‌ی میان چارکی را محاسبه کنید.

ج : نمودار جعبه ای ترسیم کنید.

۲۰ : داده های زیر را در نظر بگیرید.

۷ و ۹ و ۱۱ و ۱۲ و ۵ و ۴ و ۶ و ۱۰

الف : داده ها را طوری تغییر دهید که میانگین ۲ واحد اضافه شود ولی انحراف معیار تغییر نکند.

ب : داده ها را طوری تغییر دهید که میانگین و میانه ۳ برابر شوند.

ج : دو داده‌ی دیگر به این داده ها طوری اضافه کنید میانگین و میانه تغییر نکنند.

۲۱ : در هر مورد جای خالی را کامل کنید.

الف : تفاضل بیشترین و کمترین داده را .... گویند.

ب : در نمودار جعبه ای۵۰ درصد داده ها قبل از و۵۰ درصد داده ها بعد از .... قرار دارند.

پ : در نمودار جعبه ای یک مجموعه‌ی داده های آماری ...... درصد داده ها قبل از چارک اوّل و .... درصد داده ها قبل از چارک دوّم و .... درصد داده ها قبل از چارک سوّم قرار می گیرند.

ت : معیارهای پراکندگی که میانگین در آنها نقش دارد .... و.......... هستند.

ث : اگر داده های آماری k برابر شوند، ضریب تغییرات .... (عدد k مثبت است.)

ج : واحد اندازه گیری واریانس ، .... واحد اندازه گیری داده ها است.

ح : هر قدر انحراف معیار کمتر باشد، میزان .... داده ها، کمتر است.

***

تهیه کننده : جابرعامری عضو گروه ریاضی دوره ی دوم متوسطه استان خوزستان