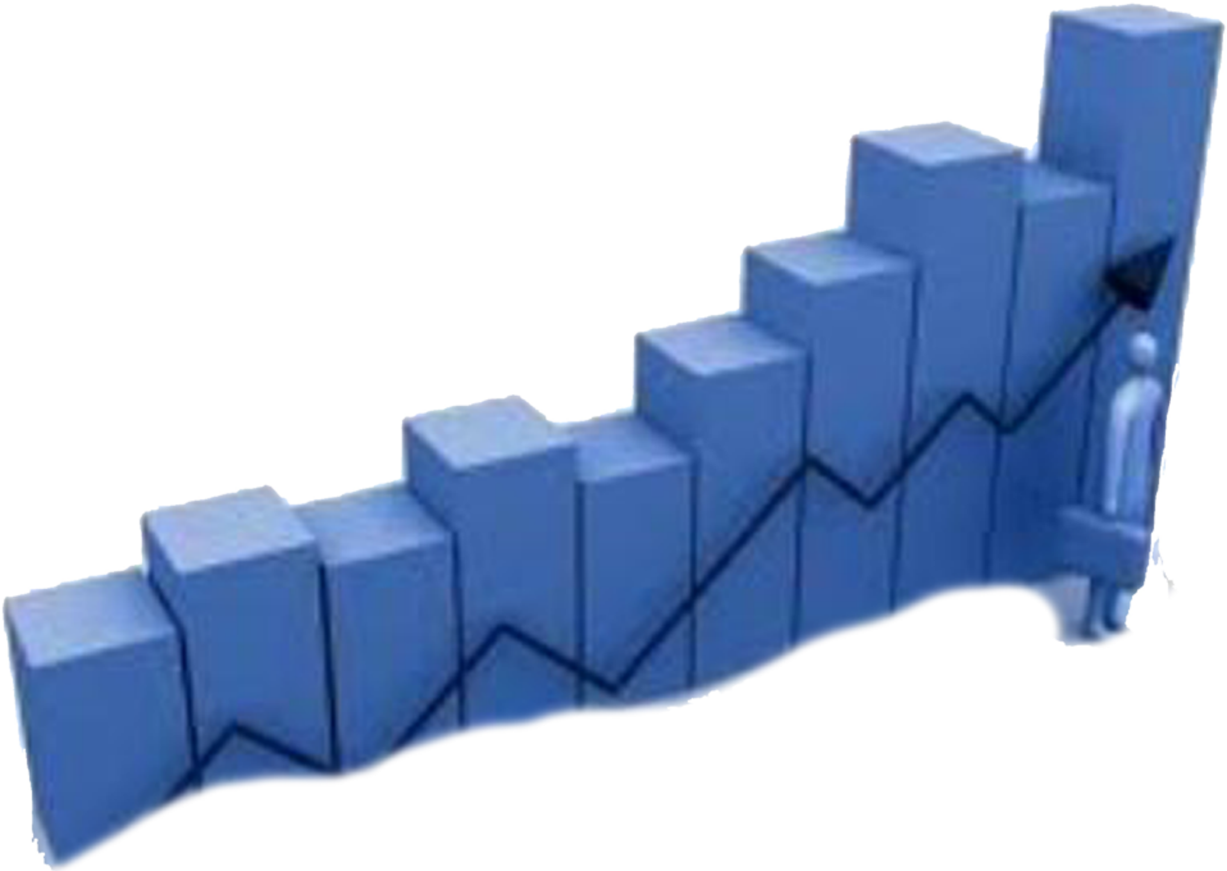
سبک جهانی کانگ فو توآ ۲۱
معاونت آموزش

آمار توصیفی



تقدیم به :
رهروان طریقت دانایی کانگ فو توآ ۲۱

گردآوری : راهدان مجتبی اسماعیلی

جامعه آماری: مجموعه ای از افراد یا اشیا است که درباره اعضای آن می خواهیم موضوع یا موضوعاتی را مطالعه کنیم

مثال: مجموعه افراد جویای کار، مجموعه پزشکان متخصص قلب و عروق، مجموعه دبیران استان کرمان

اگر تمام افراد جامعه را مورد مطالعه قرار دهیم می گوییم سرشماری کرده ایم.

مهمترین مشکلاتی که در سرشماری با آنها مواجه هستیم عبارتند از:

۱-در دسترس نبودن تمام اعضای جامعه

۲-وقت گیر بودن دسترسی به تمام اعضای جامعه

۳-گران تمام شدن بررسی تمام اعضای جامعه

۴-ازبین رفتن جامعه در برخی از مطالعات

این مشکلات سبب می شود تا سعی کنیم از راه میانبر برویم و بجای آنکه تمام اعضای جامعه را مورد مطالعه قرار دهیم بخشی از آن را که با دقت و مطالعه لازم انتخاب شده است، را بررسی کنیم. این بخش کوچک از جامعه آماری را نمونه می گوییم

نمونه: زیرمجموعه ای از جامعه است که طبق روشی خاص و با دقت انتخاب شده است جهت بررسی روی یک صفت خاص جامعه.

تعداد اعضای جامعه را اندازه جامعه می گوییم

تعداد اعضای نمونه را اندازه نمونه می گوییم

نمونه باید به قسمی انتخاب شود که بیانگر جامعه باشد . فرض کنید بخواهید درباره قد همکلاسی های خود مطالعه کنید . حال اگر افراد تیم بسکتبال کلاس را به عنوان نمونه انتخاب کرده باشید، نمونه خوبی درباره همکلاسی های خود را ارائه نکرده اید؛ زیرا همه آنها قدبلند می باشند.

روش انتخاب نمونه باید به گونه ای باشد که:

۱-امکان انتخاب هر فرد به عنوان عضوی از نمونه امکان پذیر باشد

۲-هر فرد برای شرکت در نمونه همانقدر سهم داشته باشد که دیگران دارند

داده : نتایج حاصل از اندازه گیری یا بررسی نمونه را داده می گوییم

روش جمع آوری داده ها:

۱-استفاده از داده های از پیش تهیه شده

۲- از طریق پرسش (شفاهی، مصاحبه، کتبی)

۳-از طریق مشاهده و ثبت وقایع

۴- از طریق انجام آزمایشات

متغیر تصادفی:

بعد از انتخاب نمونه به دور از نظر شخصی و سلیقه های فردی، باید موضوع یا موضوعات موردنظر را بررسی کنیم . این موضوع یا موضوعات را متغیر تصادفی می نامیم.

مثال: قد دانش آموزان یک کلاس، سابقه کار دبیران ریاضی مدرسه زینبیه، میزان د رآمد افراد شرکت مس، رنگ چشم همراهان کلاس، گروه خونی اعضای شرکت کننده در تیم والیبال شهرستان، درجه حرارت هوا در ساعت ۱۰ صبح روزهای آبان ماه سال ۹۱

انواع متغیرها:

۱ـ متغیرهای کمی: متغیرهایی هستند که قابل اندازه گیری اند.

۲- متغیرهای کیفی: متغیرهایی هستند که قابل اندازه گیری نیستند.

متغیرهای کمی:

۱ـ متغیرهای پیوسته: یک متغیر کمی است که اگر دو مقدار a و b را بتواند اختیار کند هر مقدار بین آنها را نیز بتواند اختیار کند .

مثال: وزن افراد

۲ـ متغیرهای گسسته: متغیر کمی که پیوسته نباشد، گسسته گوییم. مثال: تعداد روزهای بارانی در سال

متغیرهای کیفی:

۱ـ متغیرهای کیفی ترتیبی: متغیرهای کیفی ای که در آنها نوعی ترتیب وجود دارد را گوییم. مثال: مراحل تحصیلی شامل: دبستان–راهنمایی– دبیرستان و غیره

۲ـ متغیرهای کیفی اسمی: متغیرهایی که ترتیبی نباشند، را گوییم.

دسته بندی داده ها و جدول فراوانی:

تاکنون با جامعه، نمونه، متغیرهای تصادفی و داده ها آشنا شده ایم . پس از جمع آوری داده ها تعدادی عدد، مثلاً ۵۰ عدد در اختیار داریم. این اعداد داده های خام هستند

جدول فراوانی: معمولاً داده ها در قالب یک جدول داده م ی شوند. این داده ها به علت طبیعت ظاهری نامنظم آنها گویای مطلبی درباره جامعه نیستند. برای آنکه بتوان به آنها نظم بهتری داد جدول مناسبی تنظیم می شود

عبارات توصیفی را بهتر می توان به ذهن سپرد. از این رو برای انتقال سریع اطلاعات و نه برای منظورهای تخصصی، روش توصیفی مناسب است ولی این بیان توصیفی نمی تواند مبنایی برای تصمیم گیری های جدی شود

در هر صورت ما ناگزیر هستیم که برخی از اطلاعات را کنار بگذاریم، برخی دیگر را یک کاسه کنیم تا بتوانیم جامعه مورد مطالعه را معرفی کنیم. به همین منظور به سراغ جداول فراوانی می رویم.

فراوانی و فراوانی نسبی:

هرگاه n چیز از k نوع $T_1,T_2,T_3,T_4,\dots T_k$ به ترتیب با تعدادهای $F_1,F_2,\dots F_k$ تشکیل شده باشند، این تعدادها را فراوانی های این چیزها می گوییم و $\frac{F_1}{n},\frac{F_2}{n},\frac{F_3}{n},\dots\frac{F_k}{n}$ را فراوانی‌های نسبی چیزها می گوییم که فراوانی های نسبی را با $r_1,r_2,\dots r_k$ نشان می دهیم.

بطور ساده تعداد هر نوع را فراوانی و حاصل تقسیم فراوانی بر تعداد کل را فراوانی نسبی می گوییم.

۱–مجموع فراوانی‌ها برابر با کل داده هاست

$$\sum_{i=1}^{n} fi = n$$

۲–مجموع فراوانی‌های نسبی برابر با یک می باشد

$$\sum_{i=1}^{k} ri = n$$

فراوانی تجمی و فراوانی تجمعی نسبی

فراوانی تجمعی: اگر در هر نوع (طبقه) فراوانی آن نوع (طبقه) با فراوانی ماقبل خود جمع شود فراوانی تجمعی بدست می آید.

فراوانی تجمعی نسبی: اگر در هر نوع (طبقه) فراوانی نسبی آن طبقه و فراوانی نسبی طبقات قبل از آن با هم جمع کنیم فراوانی تجمعی نسبی بدست می آید.

J= شماره طبقات

فراوانی تجمعی نسبی = $g_i = f_1 + f_2 + f_3 + \cdots f_j = \sum_{i=1}^{j} f_i$

که فراوانی تجمعی طبقه آخر برابر n می باشد $g_k=n$

فراوانی تجمعی نسبی طبقه آخر برابر با یک می باشد $s_k=1$

مثال (۱)

پزشکی ۲۰ بیمار قلبی دارد که گروه خونی آنها عبارتند از:

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

جدول فراوانی را رسم می کنیم

ابتدا باید بدانیم که نوع متغیر در اینجا متغیر اسمی می باشد

Xi = نماینده هر طبقه یا گروه که با یک نشان داده شده است.

Fi = فراوانی یا تعداد هر گروه خونی

Ri =فراوانی نسبی آن گروه می باشد

g_i = فراوانی تجمعی

S_i = فراوانی تجمعی نسبی هر گروه می باشد.

گروه خونی	x_i	f_i	r_i	g_i	S_i
A	۱	۶	٪۳۰	۶	٪۳۰
B	۲	۳	٪۱۵	۹	٪۴۵
AB	۳	۴	٪۲۰	۱۳	٪۶۵
O	۴	۷	٪۳۵	۲۰	۱/۰۰
		۲۰	۱/۰۰		

گروه A $r_1 = \frac{6}{20} = \frac{3}{10} = \%30$ حاصل $\frac{f_i}{n}$

گروه B $r_2 = \frac{3}{20} = \%15$

گروه AB $g_3 = f_1 + f_2 + f_3 = 6 + 3 + 4 = 13$

گروه AB $s_3 = r_1 + r_2 + r_3 = \%30 + \%15 + \%20 = \%65$

جدول فراوانی داده های پیوسته

برای رسم جدول برای داده های پیوسته به نکات زیر دقت فرمایید.

۱ـ نماینده هر طبقه در داده های پیوسته برابر است با نقطه میانگین آن رده یعنی

$$x_i = \frac{(\text{کران بالا} + \text{کران پایین})}{2}$$

۲ـ برای کران بالا و پایین، ابتدا بزرگترین و کوچکترین داده ها را پیدا نموده و از کوچکترین به اندازه ۰/۵ کم و به بز رگترین ۰/۵ اضافه می کنیم و این به این دلیل است که دو عدد کوچکتر و بزرگتر هم در داده ها شمارش شوند

بقیه موارد جدول فراوانی مشابه می باشد با جدول فراوانی داده های گسسته

۳ـ معمولاً تعداد رده ها را بین ۵ تا ۲۵ انتخاب می نمایند، یک قاعده مفید استفاده از دستور STUR GES می باشد که در آن تعداد رده ها K از رابطه زیر بدست می آید

تعداد رده ها (طبقات) $K = 1/33\log\frac{n}{10}$

n: برابر کل داده هاست

K: تعداد طبقات می باشد

نکته: عدد بدست آمده اعشاری می باشد که حاصل را به عدد بالاتر گرد می کنیم

مثال:

داده های زیر اندازه های قد ۱۰۰ جوان بیست ساله در یکی از شهرستانهای ایران می باشند که برحسب سانتیمتر تا نزدیکترین واحد سرراست شده اند.

این داده ها نتیجه اندازه گیری با مقیاس نسبتی هستند و آنها را داده های پیوسته می گویند برای تشکیل جدول فراوانی به طریق زیر عمل می کنیم

–در این داده ها عدد ۱۵۰ کوچکترین و عدد ۱۸۴ بزرگترین داده هستند.

از کمترین داده ها به اندازه ۰/۵ کم و به بزرگترین داده ۰/۵ اضافه می کنیم پس می توان گفت اندازه واقعی داده ها (قدها) در فاصله [۱۸۴/۵ و ۱۴۹/۵] قرار دارند. طول این فاصله یعنی ۳۵=۱۴۹/۵− ۱۸۴/۵ را بُرد داده ها می نامیم. و مثلاً به ۵ قسمت مساوی تقسیم می کنیم ۷=۵÷۳۵ پس طول هر رده برابر ۷ می شود.

رده اول ۱۵۶/۵=۷+ ۱۴۹ /۵ ⟵ [۱۵۶/۵ و ۱۴۹/۵] رده اول با طول واقعی ۷ می باشد

رده دوم ۱۶۳/۵=۷+۱۵۶/۵ ⟵ [۱۶۳/۵ و ۱۵۶/۵] رده دوم با طول واقعی ۷ می باشد.

رده سوم ۱۷۰/۵=۷+۱۶۳/۵ ⟵ [۱۷۰/۵ و ۱۶۳/۵] رده سوم با طول واقعی ۷ می باشد.

ـ نقطه وسط هر رده که آن را با X_i نشان می دهیم نماینده آن رده می باشد که برابر است با $x_i = \frac{\text{کران پایین} + \text{کران بالا}}{2}$

نماینده طبقه اول $\frac{149/5 + 156/5}{2} = 153$

نماینده طبقه دوم $\frac{156/5+163/5}{2}=160$

نماینده طبقه پنجم $\frac{177/5+184/5}{2}=181$

جدول فراوانی قد ۱۰۰ جوان بیست ساله با متغیر نسبتی X_i

رده	x_i	f_i	r_i	g_i	S_i
۱۴۹/۵ – ۱۵۶/۵	۱۵۳	۱۵	$r_1=\frac{15}{100}=\%15$	۱۵	$\frac{15}{100}=\%15$
۱۵۶/۵ – ۱۶۳/۵	۱۶۰	۲۰	$r_2=\frac{20}{100}=\%20$	۱۵+۲۰=۳۵	$\frac{35}{100}=\%35$
۱۶۳/۵ – ۱۷۰/۵	۱۶۷	۳۰	$\frac{30}{100}=\%30$	۳۰+۳۵=۶۵	$\frac{65}{100}=\%65$
۱۷۰/۵ – ۱۷۷/۵	۱۷۴	۲۵	$\frac{25}{100}=\%25$	۲۵+۶۵=۹۰	$\frac{90}{100}=\%90$
۱۷۷/۵ – ۱۸۴/۵	۱۸۱	۱۰	$\frac{10}{100}=\%10$	۱۰+۹۰=۱۰۰	$\frac{100}{100}=1$
کل		۱۰۰	۱۰۰		

$$r_i=\frac{f_i}{n} \qquad r_1=\frac{f_1}{n}=\frac{15}{100}=\%15$$

$$r_5=\frac{f_5}{n}=\frac{10}{100}=\%10$$

$$g_1=15 \qquad g_2=f_1+f_2=15+20=35 \qquad g_3=f_1+f_2+f_3=15+20+30=65$$

$$s_i=\frac{g_i}{n} \qquad s_1=\frac{g_1}{n}=\frac{15}{100}=\%15 \qquad s_2=\frac{g_2}{100}=\frac{35}{100}=\%35$$

انواع نمودارها:

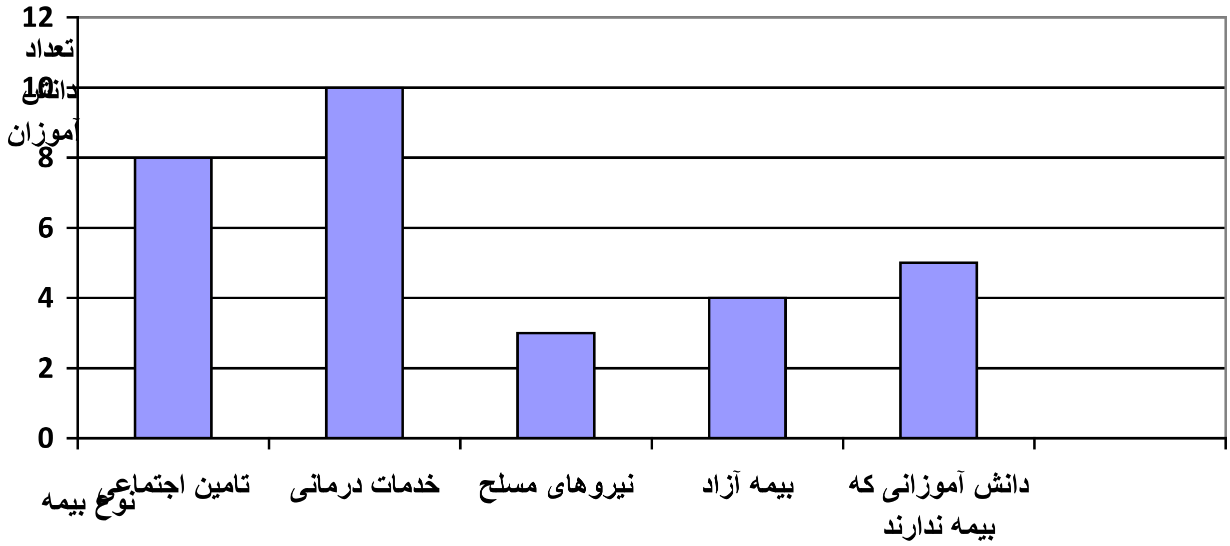
نمودار چیست؟ نمایش داده ها را طبق قراردادهای خاص به صورت هندسی، یک نمودار آماری گویند

نمودار میله ای: این نمودار بیشتر برای متغیرهای گسسته و کیفی مناسب است.

مثال: وضعیت بیمه دانش آموزان کلاس ۳۰ نفره به صورت زیر می باشد نمودار میله ای آن را رسم کنید

برای رسم نمودار فراوانی ها را در محور عرض (ارتفاع) قرار می دهیم و وضعیت بیمه را در محور طول قرار می دهیم

وضعیت بیمه	تعداد (فراوانی)
بیمه تأمین اجتماعی	۸
بیمه خدمات درمانی	۱۰
بیمه نیروهای مسلح	۳
شرکت بیمه آزاد	۴
دانش آموزانی که بیمه ندارند	۵



نمودار دایره ای: این نمودار بیشتر برای داده های گسسته کیفی که دارای حالتهای مختلف باشد بکار می رود

مثلاً مقطع تحصیلی، متغیر تصادفی کیفی دارای ۴ حالت ابتدایی، راهنمایی، دبیرستان و دانشگاه می باشد

فرض کنید فراوانی حالت اول F_1 و حالت دوم F_2 و الی آخر...... F_n باشد.

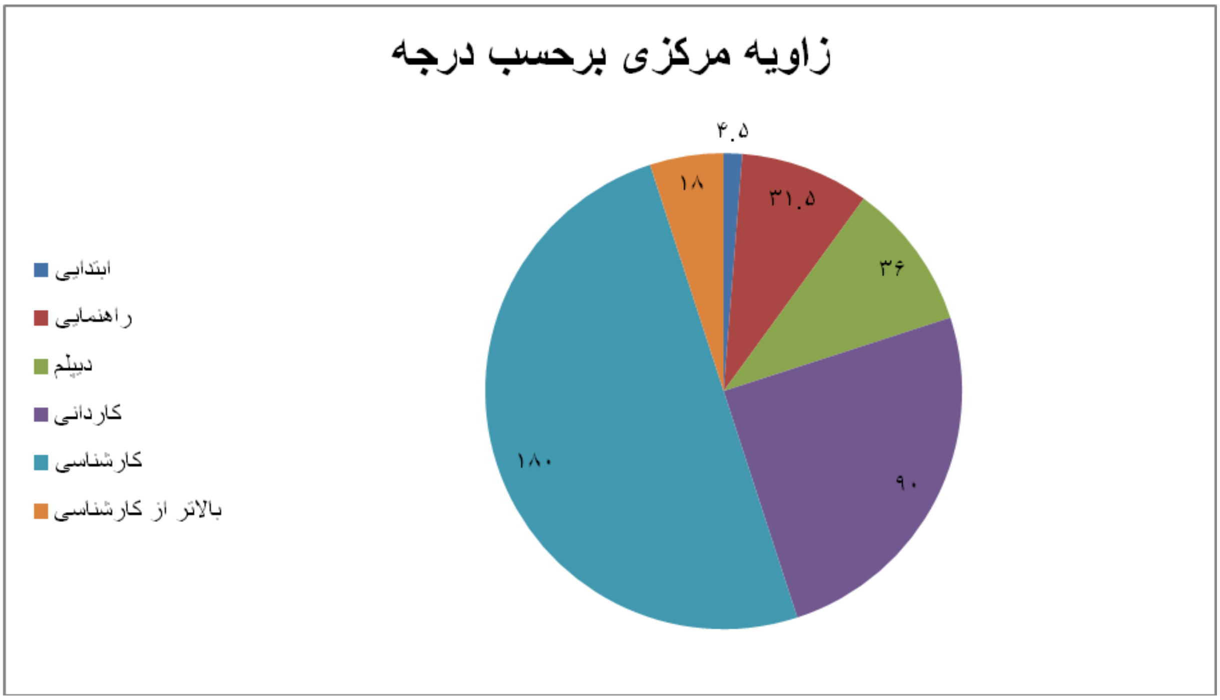
زاویه مرکزی نظیر هر دسته یا حالت برابر است با $\frac{f_i}{n} \times 360 =$ زاویه مرکزی هر داده

مثال: یک مؤسسه بزرگ ۴۰۰ نفر کارمند دارد که برحسب سطح سواد به صورت زیر توزیع شده اند نمودار دایره ای آن را رسم کنید.

سطح سواد	ابتدایی	راهنمایی	دیپلم	کاردانی	کارشناسی	بالاتر از کارشناسی
تعداد کارمندان	۵	۳۵	۴۰	۱۰۰	۲۰۰	۲۰

زاویه ها مطابق فرمول به صورت زیر محاسبه می شوند

حالت	فراوانی	زاویه مرکزی برحسب درجه
ابتدایی	۵	$\frac{360}{400} \times 5 = 0/9 \times 5 = 4/5^\circ$
راهنمایی	۳۵	$0/9 \times 35 = 31/5^\circ$
دیپلم	۴۰	$0/9 \times 40 = 36^\circ$
کاردانی	۱۰۰	$0/9 \times 100 = 90^\circ$
کارشناسی	۲۰۰	$0/9 \times 200 = 180^\circ$
بالاتر از کارشناسی	۲۰	$0/9 \times 20 = 18^\circ$



نمودار مستطیلی: این نمودار برای متغیرهای کمی پیوسته مناسب است.

نموداری است مرکب از چند مستطیل که از روی جدول فراوانی داده های پیوسته ساخته می شود.

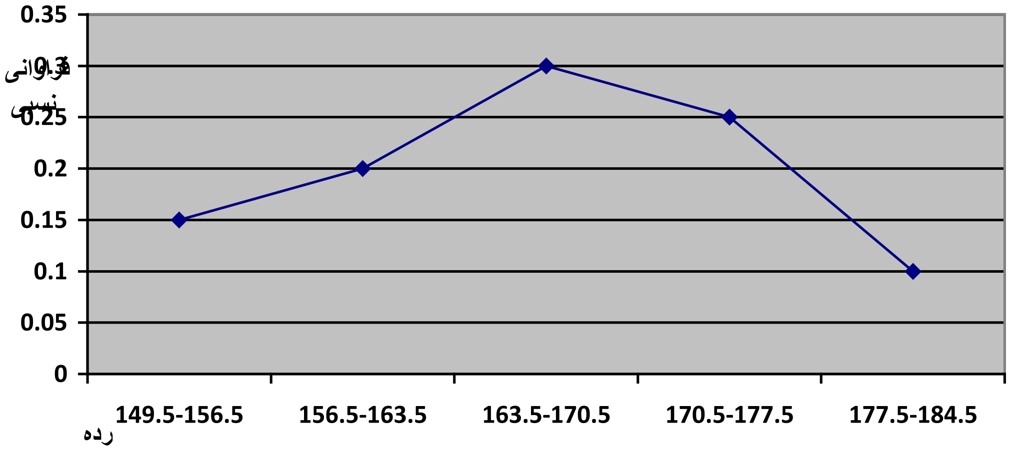
تعداد مستطیل ها برابر تعداد رده ها و قاعده هر مستطیل روی محور x ها جا دارد و طولش برابر است با طول واقعی رده که هرچه باشد آن را یک واحد تلقی می کنیم.

مثال: نمودار ستونی را برای جدول فراوانی قد بیست ساله رسم می نماییم. هم می توانیم از فراوانی استفاده کنیم هم از فراوانی نسبی در اینجا به ستون رده و فراوانی نسبی نیاز داریم.

ـ نمودار فراوانی نسبی داده های قد ۱۰۰ جوان ۲۰ ساله

ـ نمودار چندبر فراوانی نسبی داده های قد ۱۰۰ جوان ۲۰ ساله

رده	r_i
۱۴۹/۵ –۱۵۶/۵	٪۱۵
۱۵۶/ ۵–۱۶۳/۵	٪۲۰
۱۶۳/۵ –۱۷۰/۵	٪۳۰
۱۷۰/۵ –۱۷۷/۵	٪۲۵
۱۷۷/۵ ـ۱۸۴/۵	٪۱۰
کل	٪۱۰۰



حال اگر در حین نمودار فراوانی نسبی نقطه وسط هر طبقه با همان فراوانی نسبی درنظر بگیریم و سپس این نقاط را بهم وصل کنیم نمودار چندبر فراوانی بدست می آید.

نمودار چندبر فراوانی که در صفحه قبل توضیح داده شد برای داده های پیوسته رسم می شود

شاخصهای مرکزی

۱ـ مُد: داده ای است که بیشترین فراوانی را دارد.

ممکن است مد منحصر به فرد نباشد و نمونه ای دارای ۲ مدی یا چند مدی باشند.

در رأی گیری ها اساس تصمیم گیری مُد است.

مد در داده های زیر کدام است؟

۵ و ۸ و ۷ و ۵ و ۴ و ۵ و ۵ و ۲ و ۱ و ۴ و ۳ و ۲ و ۱ و ۱

در داده های بالا مد برابر M=۵ می باشد.

۲ـ میانه:

در داده های گسسته : ابتدا داده ها را از کوچک به بزرگ مرتب می کنیم در صورتی که تعداد داده ها فرد باشد . داده ای را که در وسط قرار دارد، میانه می گویند.

در صورتی که داده ها زوج باشند مجموع دو داده وسط تقسیم بر ۲ میانه می شود.

مثال: میانه را برای داده های زیر پیدا نمایید

الف) ۳۱ , ۴ ,۵ ,۲۱ , ۲۰, ۳۰ ,۱۷ , ۸ , ۱۲

ب) ۲ , ۳ , ۲ ,۱ , ۴ , ۵ , ۳ , ۲

ابتدا به ترتیب داده ها را از کوچک به بزرگ مرتب می کنیم

داده های فرد — الف) m=۱۷

داده های زوج

$$m = \frac{2+3}{2} = 2/5$$

ب)

میانگین: شما تاکنون بارها معدل خود را حساب کرده اید برای این کار جمع نمرات را بر تعداد آنها تقسیم کرده اید

درواقع میانگین یعنی معدل داده ها

حال اگر داده ای چند بار تکرار شود یعنی دارای فراوانی بیش از یک باشد به جای اینکه آن را چند بار جمع بزنیم در فراوانی اش ضرب می کنیم و بر تعداد داده ها تقسیم می کنیم.

اگر K طبقه داشته باشیم و هر کدام دارای فراوانی f_i باشن میانگین به صورت زیر محاسبه می شود

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i \times i}{n}$$

مثال: اعداد زیر نمره های ریاضی یک کلاس می باشد میانگین آنها را محاسبه نمایید

۱۰۰ , ۱۰۰ ,۷۰ , ۶۳ , ۶۲ , ۱۰ , ۱۲ , ۶۰

$$\bar{X} = \frac{(100 \times 2) + (70) + (63) + (62) + (10) + (12) + (60)}{8} = 47$$

میانگین را برای داده های زیر محاسبه نمایید

ستون $f_i x_i$ را تشکیل می دهیم.

X_i	F_i	$F_i x_i$
۲	۸	۲×۸=۱۶
۳	۱۰	۳×۱۰=۳۰
۴	۱۲	۴×۱۲=۴۸
۵	۷	۵×۷=۳۵
۶	۵	۶×۵=۳۰
۷	۴	۷×۴=۲۸
۸	۴	۸×۴=۳۲
	۵۰	۲۱۹

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i x_i}{n} = \frac{219}{50} = 4/38$$

شاخصهای پراکندگی:

۱)دامنه تغییرات: برابر است با اختلاف بزرگترین داده و کوچکترین داده

R=max-min

مثال دامنه تغییرات نمرات درس فیزیک دانش آموزان یک مدرسه که به صورت زیر می باشند را محاسبه نمایید

۱۰ , ۱۱ , ۱۲ , ۱۳ , ۱۴ , ۱۵ , ۱۶ , ۱۷ , ۱۸ , ۱۹

R=۱۹-۱۰=۹

۲)واریانس:

واریانس برابر است با میانگین مجذور انحرافات از میانگین است و آن را با δ^2 نشان می دهیم.

بنابراین

$$\delta^2 = \frac{(x_1-\bar{x})^2+\cdots(x_n-\bar{x})^2}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

[δ از حروف کوچک یونانی است و سیگما خوانده می شود]

مثال: تعداد ساعاتی که ۴ دانش آموز در طول هفته به ورزش اختصاص داده اند در زیر آمده است:

واریانس این داده ها را حساب کنید.

۱ و ۵ و ۷ و ۹

میانگین = $\bar{X} = \frac{1+5+7+9}{4} = \frac{22}{4} = 5/5$

۳)انحراف معیار:

این واحد واریانس از نوع مجذور واحد متغیر است می تواند مشکلات و سوء تفاهماتی را درپی داشته باشد . فرض کنید شما قدها را برحسب متر اندازه گیری کرده باشید و همکلاسی شما قده ا را برحسب سانتیمتر، در این صورت واریانسی که همکلاسی شما حساب کرده است ۱۰۰۰۰ برابر واریانسی است که شما حساب کرده اید. همکلاسی شما ممکن است اینطور نتیجه بگیرد که پراکندگی در قد بسیار زیاد و چشمگیر است که غیرممکن است . برای اینکه اختلاف نظرها را ازبین ببریم س‌عی می کنیم که تفاوت عمده در واحد واریانس و واحد میانگین را با جذر گرفتن از واریانس ازبین ببریم

انحراف معیار: که با نماد δ نشان داده می شود برابر جذر واریانس است.

مثال: تعداد ساعتهایی که در هفته ۴ نفر به صورت داوطلبانه در یک بیمارستان خدمت می کنند در زیر آمده است انحراف معیار داده ها را حساب کنید.

۱ و ۳ و ۴ و ۵ و ۶

$$\bar{X} = \frac{1+3+4+5+6}{5} = \frac{19}{5} = 3/8$$

$$\delta^2 = \frac{(1-3/8)^2+(3-3/8)^2+(4-3/8)^2+(5-3/8)^2+(6-3/8)^2}{5}$$

$$\delta^2 = \frac{(-2/8)^2+(-0/8)^2+(-0/2)^2+(-1/2)^2+(-2/2)^2}{5}$$

$$= \frac{7/84+0/64+0/04+1/44+4/84}{5}$$

$$\delta^2 = \frac{14/80}{5} = 3/96$$

$$\delta = \sqrt{3/96} \approx 1/7$$