ریاضی عمومی و مقدمات آمار

(ویژه دانشجویان تربیت بدنی)

تالیف: دکتر سیدحسین علوی

استادیار دانشکده فنی و حرفه ای - دانشگاه فنی و حرفه ای

انتشارات نوروزی

فصل 1

توزیع فراوانی

جدول توزیع فراوانی

پس از جمع آوری اطلاعات و دست یافتن به نتایج، پژوهشگر با این مسئله روبه روست که با نتایج حاصله چه نوع فعالیت یا فعالیت‌های انجام دهد؟ اولین کاری که تصویر روشن‌تری از نمره‌های یک کلاس بدست خواهد داد این است که نمره‌ها را به ترتیب از کوچک به بزرگ(نمره کوچک از پایین) تنظیم کند.

مراحل تشکیل جدول توزیع فراوانی

دامنه تغییرات

دامنهٔ تغییرات یعنی فاصلهٔ بین بزرگترین و کوچکترین عدد. اگر بزرگترین را با H و کوچکترین داده رابا L و دامنهٔ تغییرات را با R نشان دهیم، دامنه‌ی تغییرات از فرمول زیر به‌دست می‌آید:

$R = (H - L) + \alpha$ یا $R = (X_{max} - X_{min}) + \alpha$

مقدار α برای اعداد اعشاری که بخش اعشاری یک رقمی است 0/1 می‌باشد و برای حالتی که بخش اعشاری دو رقمی است 0/01 می‌باشد و برای سایر اعداد اعشاری از فرآیند مشابهی استفاده می‌شود. مقدار α برای اعداد صحیح برابر یک است.

تذکر: در بعضی کتاب‌ها α را در نظر نمی‌گیرند و دامنه تغییرات را از فرمول زیر به‌دست می‌آورند.

$$R = (X_{max} - X_{min})$$

اگردامنهٔ تغییرات کوچکتر از20 باشد نیازی به طبقه‌بندی اعداد نیست و از جدول توزیع فراوانی ساده است استفاده می‌کنیم.

مثال: اگر بزرگترین عدد یک توزیع 25 و کوچکترین عدد آن 15 باشد دامنه تغییرات به‌صورت زیر محاسبه می‌شود.

حل:

$$R = (H - L) + 1$$

$R = (25-15)+1 \Rightarrow R = 11$

مثال: اگر بزرگترین عدد یک توزیع 8 /10 و کوچکترین عدد آن 6 /7 باشد دامنه تغییرات آن به صورت محاسبه می‌شود.

$R = (10/8 - 7/ 6) +0/1 \Rightarrow R =3/3$

همان‌طور که قبلاً توضیح داده‌ایم برای اعداد اعشاری یک رقمی، α برابر 1/ 0 است.

مثال: اگربزرگترین عدد یک توزیع 25 /20 وکوچکترین عدد18/21 باشد دامنه تغییـرات آن بـه صورت زیر محاسبه می‌شود.

$R = (20 / 25 - 18 / 21) + 0/01 \Rightarrow R = 2/ 0 5$

توجه کنید که در اینجا اعداد اعشاری دو رقمی داریم و لذا α برابر یک صدم است.

مثال : دامنه تغیرات 7/3، 6/2، 8/1.11/12 ، 12/7 و 7/1 را محاسبه کنید.

حل:

$R = (H - L) + \alpha$

$R = (12/7 - 6/2) + 0/01 = 6/5 + 0/01 = 6/51$

چون در نمرات دو اعشاری وجود دارد، پس $\alpha = 0/01$ می‌شود.

جدول توزیع فراوانی ساده

برای تشکیل جدول توزیع فراوانی ابتدا دامنه تغییرات را تشکیل می‌دهیم.

اگر دامنه‌ی تغییرات به دست آمده کوچکتر از 20 باشد، از جدول توزیع فراوانی ساده اسـتفاده می‌کنیم. برای تشکیل جدول توزیع فراوانی، ابتدا جدولی با 3 ستون رسم مـی‌کنـیم و اعـداد را بـه ترتیب از بالا به پایین و از بزرگ به کوچک می نویسیم. سپس به ازای هر مرتبۀ تکرار هر عـدد، در ستون دوم جلوی آن یک خط نشان و یا علامت(/) قرار می دهیم و درآخر، تعداد خط و نشان‌ها را به عنوان فراوانی مطلق یا (f) در ستون سوم می‌نویسیم. در واقع فراوانی مطلـق یـک داده آمـاری، تعداد دفعاتی است که آن داده تکرار شده است.

مثال: اعداد زیر نتیجه تست دراز و نشست 28 دانش‌آموز مـی‌باشـد. جـدول توزیـع فراوانـی آن را تشکیل دهید.

29, 34, 31, 42, 34, 30, 41, 32
35, 40, 32, 35, 35, 40, 35, 33
29, 30, 33, 35, 34, 32 , 42
35, 34, 30, 31, 35

حل:

جهت تشکیل جدول توزیع فراوانی ابتدا دامنه‌ی تغییرات را محاسبه می‌کنیم. در اعداد بالا، بزرگترین عدد 42 و کوچکترین عدد 29 می‌باشد. دامنه تغییرات آن را محاسبه می‌کنیم.

$$R = (H - L)+1 \Rightarrow R=(42 - 29) +1 \; R= \Rightarrow 13 + 1 \; R= \Rightarrow 14$$

چون دامنه تغییرات بدست آمده کمتر از 20 است، احتیاجی به طبقه‌بندی نیست. بنابراین، داریم:

جدول 1-4. توزیع فراوانی ساده

X	خط و نشان	f
42	//	2
41	/	1
40	//	2
35	~~////~~ //	7
34	////	4
33	//	2
32	///	3
31	//	2
30	///	3
29	//	2

جدول توزیع فراوانی طبقه بندی شده

اگر دامنه تغییرات به دست آمده از 20 بیشتر شد، برای سهولت در محاسبات، جدول طبقه بندی شده تشکیل می‌دهیم. برای تشکیل جدول طبقه بندی مراحل زیر انجام می دهیم.

محاسبه دامنه تغییرات: همان‌طوری‌که که قبلاً توضیح داده‌ایم دامنه تغییرات از فرمول $R = (H - L) + \alpha$ به‌دست می‌آید.

تعداد طبقات:

برای تعداد طبقات می‌توان از فرمول $K = 1 + 3/3 \; Log_n$، که N تعداد اعداد و K تعداد طبقات است، استفاده کرد.

مثال: چنانچه تعداد اعداد جمع آوری شده برابر 60 باشد، تعداد طبقات را محاسبه کنید .

حل:

پایه لگاریتم بر مبنای 10 است .

$$K = 1 + 3/3 \log n \qquad k = 1 + 3/3 \log 60$$

$\text{Log}\,60$ در پایه 10 مساوی 1/7782 است.

$K = 1 + 3/3(1/7782) = 6/86806 \approx 7$

فاصله طبقه‌ای: در انتخاب فاصله طبقاتی می‌توان به دو نکته توجه کرد. نخست آن که تعداد طبقه‌ها کمتر از 10 و بیشتر از 25 نباشد. اگر تعداد طبقه‌ها کمتر از 10 باشد، اطلاعات زیادی به خاطر طبقه‌بندی از دست می‌رود و در نتیجه از دقت محاسبات کاسته می‌شود؛ و اگر تعداد آنها از 25 تجاوز کند انجام محاسبه دشوار می‌گردد. دوم آن که تجربه نشان داده است که برخی از اعداد مانند 1 ، 2 ، 3 ، 5 ،10 برای اندازه فاصله طبقه مناسب‌تر است. برای تعیین اندازه فاصله طبقه‌ای از فرمول زیر استفاده می‌شود.

$$\text{فاصله طبقات} = \frac{\text{دامنه تغییرات}}{\text{تعداد طبقات}} \Rightarrow I = \frac{R}{k}$$

نکته: دسته بندی باید به گونه‌ای انجام شود که فراوانی هیچ طبقه‌ای صفر نشود.

نکته:با افزایش تعداد طبقات باعث می‌شود فاصله طبقات کمتر گردیده و این اقدام موجبات کاهش خطای گروه‌بندی را فراهم می‌سازد.

تعیین پایین‌ترین طبقه: باید حد نشانی پایین‌ترین نمره بر فاصله طبقه قابل تقسیم و نتیجه آن عدد صحیح باشد.(یعنی باید کوچکترین عدد توزیع یا کوچکتر از آن که بر فاصله طبقه قابل تقسیم باشد را انتخاب کرد) طبقه a تا b شامل اعداد $a+1, a+2, a+3+a+4, \ldots, b$ است که با a-b نمایش داده می شود. پس طبقه 34 ـ 30 طبقه‌ای با فاصله طبقه‌ای 5 است که اعداد 30 ، 31 ،32 ،33 ،34 در آن قرار می‌گیرند.

نقطه میانی: حاصل جمع دو عدد نشانی طبقه تقسیم بر 2 را گویند.

مثال: اگر فاصله طبقاتی 34 -30 باشد. نقطه میانی برابر 32 = $\frac{30+34}{2}$ است.

مثال: نمرات زیر از آزمون پرش زیگزاگ به دست آمده است.جدول توزیع فراوانی‌آن را تشکیل دهید.

$49-50-61-65-72-78-36-69-80-64-75-81-78-69$

$78-90-45-52-62-58-86-72-60-94-73-81-77-56-70$

$40-58-66-61-47-63-85-53-83-47-89-73-65-56-48$

$66-73-69-41-66-60$

حل:

ابتدا دامنه تغییرات را محاسبه می‌کنیم:

$R= (H-L)+1$

$R = (94 - 36) + 1 \quad R= \Rightarrow 59$

بعد از محاسبه دامنه تغییرات تعداد طبقات را مشخص می‌کنیم، با توجه به تعداد نمرات حدود 12 طبقه در نظر می‌گیریم.

تعیین فاصله طبقه‌ای: اگر دامنه تغییرات را تقسیم بر تعداد طبقات کنیم فاصله طبقات بدست می‌آید.

$$I = \frac{R}{K} \Rightarrow I = \frac{59}{12} \approx 4/9 \Rightarrow I = 4/9$$

تذکر: فاصله طبقاتی می‌تواند اعشاری باشد ولی برای سهولت انجام کار I را مساوی 5 می‌گیریم.

چون فاصله طبقات برابر پنج شد، بهتر است طبقات را از35 که ضریبی از عدد پنج است شروع کنیم. توجه کنید عدد 35 از کوچکترین داده ما یعنی 36 کوچکتر است. طبقهٔ اول را 39 – 35 در نظر گرفته و طبقات بعدی را بر اساس آن مشخص می‌کنیم.

نکته: حد بالای اولین طبقه نباید از بزرگترین عدد، کوچکتر باشد و حد پایین آخرین طبقه ، نباید بزرگتر از کوچکترین داده باشد.

برای به‌دست آوردن عدد میانی طبقات، مجموع بزرگترین عدد و کوچکترین عدد هرطبقه را بر 2 تقسیم می‌کنیم. مثلاً برای به‌دست آوردن عدد میانی طبقه اول چنین عمل می‌کنیم.

$35 + 39 = 74 \Rightarrow 74 \div 2 = 37$ عدد میانی

نتایج حاصل از طبقه‌بندی در جدول زیر خلاصه می‌شود.

جدول 1: توزیع فراوانی طبقه‌بندی شده

طبقات	خط و نشان	فراوانی	عدد میانی
90-94	//	2	$(\frac{90+94}{2})=92$
85-89	///	3	87
80-84	////	4	82
75-79	~~////~~	5	77
70-74	~~////~~ /	6	72
65-69	~~////~~ ///	8	67
60-64	~~////~~ //	7	62
55-59	////	4	57
50-54	///	3	52
45-49	~~////~~	5	47
40-44	//	2	42
35-39	/	1	37

فراوانی نسبی (p) و روش محاسبه آن

فراوانی مطلق، معیاری مناسبی برای توصیف داده‌ها در جامعه نمی‌باشد به‌عنوان مثال فرض کنیدکه دروازه‌بانی در 5 ضربه پناتی 5 گل بخورد، و دروازه‌بان دیگری در 10 ضربه پنالتی 5 گل بخورد، در این صورت فراوانی مطلق تعداد گل‌های خورده برای هر دو دروازه بان برابر 5 است که معیار مناسبی برای بررسی وضعیت آن‌ها نیست. بهتر است که تعداد گل‌های خورده را به نسبت تعداد ضربات پنالتی برای هر فرد در نظر بگیریم. **در حالت کلی چنانچه فراوانی مطلق را بر مجموع فراوانی‌های مطلق تقسیم کنیم، فراوانی نسبی[1] آن طبقه به‌دست می‌آید**؛ که نشان دهنده این است که هر یک از داده‌ها چه درصدی از کل داده‌ها را تشکیل می‌دهند.

$$\text{فراوانی نسبی} = \frac{\text{فراوانی مطلق هر طبقه}}{\text{فراوانی کل}}$$

[1] . Relative Frequency

فراوانی مطلق تجمعی (تراکمی)

برای محاسبه فراوانی تجمعی[1] کافی است که فراوانی‌های مطلق را به ترتیب از پایین به طبقـه بـالا جمع کنیم. این نوع فراوانی‌ها به ما نشان می‌هد چه تعداد نمره پایین‌تر از حد واقعی بـالای طبقـه قرار گرفته است به طور مثال، اگر فراوانی تراکمی نمره 16 برابر 20 باشد، به ما نشان می‌دهـد کـه 20 نفر نمره‌ی برابر 16 و یا کمتر از 16 گرفتند.

نکته: فراوانی تجمعی بالاترین دسته با فراوانی کل برابر است.

نکته: در سطح اندازه‌گیری اسمی، هرگز نمی‌توان به محاسبه فراوانی تراکمی مبادرت ورزید.

فراوانی های نسبی تراکمی (CP)

برای محاسبه فراوانی نسبی تراکمی[2]، فراوانی نسبی را به ترتیب از طبقه پایین به طبقـه بـالا جمع می‌کنیم. این فراوانی به ما نشان می‌دهد چند درصد از نفرات کمتر از حد واقعی بـالایی یـک نمره و یا یک طبقه معین در کلاس قرار دارند. به طور مثال، اگر فراوانی نسبی تراکمی طبقه 62–60 برابر 39 درصد باشد، نشان‌دهنده این است که 39 درصد نفرات کلاس نمره مسـاوی و یـا کمتر از آن گرفته‌اند. یا طبقه 62-60 از 39 درصد افراد نمره بیشتر گرفته است.

---

[1] . Cumulative Ferquency

[2] . Relative Frequency - Cumulative

جدول. توزیع فروانی‌های نمرات خام مربوط به 150 دانش آموز هنرستان شهید کلانتری محمودآباد در پرش ارتفاع جفتی که طبق نورم‌های استاندارد بین صفر تا صد به آنها نمره داده شده است، به صورت زیر است.

79	51	67	50	78	71	77	75	55	65
62	89	83	73	80	67	74	63	32	88
88	48	60	71	79	79	47	55	70	34
89	81	46	50	61	72	86	68	75	93
41	50	90	75	61	82	73	66	54	58
59	37	42	72	80	64	67	57	87	41
75	73	79	67	74	78	91	51	36	52
70	73	77	36	85	74	70	69	95	76
67	67	85	74	77	45	39	62	76	69
91	43	42	93	83	78	73	59	53	57
93	63	55	79	71	81	84	61	47	43
71	98	53	96	77	83	72	38	73	57
70	92	59	86	53	71	49	67	81	75
33	67	67	71	71	59	80	68	42	46
82	68	30	72	57	52	50	66	39	49

تعداد طبقات: باتوجه به تعداد نمرات خام که 150 عدد است حدود 15 طبقه پیش‌بینی می‌شود.

تعیین فاصله طبقه ای: ($I=$ ؟)

ابتدا دامنه‌ی تغییرات را محاسبه می‌کنیم:

$R=(H-L)+1$

$R = (98 - 30) + 1 = 68 + 1 \quad R \Rightarrow 69$

$$I = \frac{R}{K} = \frac{69}{15} \approx 4/67 \Rightarrow i = 5$$

فاصله طبقات = Interval و دامنه طبقات = Range

چون فاصله طبقات را گرد کرده‌ایم برای آن که فراوانی هیچ طبقه‌ای صفر نشود، 14 طبقـه در نظر می‌گیریم.

تعیین پایین‌ترین نقطه: کوچکترین عدد 30 است که اگر بر فاصـله طبقـاتی تقسـیم شـود، حاصل آن عددی صحیح و زوج می‌شود ($6 = 5 \div 30$). بنابراین با توجه به فاصله طبقاتی 5 ، حد نشانی کوچکترین عدد که 30 است اولین طبقه 34 – 30 تعیین مـی گـردد و طبقـات بعـدی بـه ترتیب بالای آن قرار می‌گیرد. نتایج در جدول زیر خلاصه شده است .

جدول . انواع فراوانی‌ها را در نمره های دسته بندی شده را نشان می‌دهد.

حدهای واقعی‌طبقه‌ها	نشان دسته	فراوانی‌تجمعی نسبی (cp%)	فراوانی‌نسبی (P%)	فراوانی‌تجمعی مطلق (cf)	فراوانی‌مطلق (f)	حدود طبقات
94/5–99/ 5	97	1 /00	0/ 0 2	150	3	95-99
89/5–94/ 5	92	0/ 98	0/ 05	147	7	90-94
84/5–89/ 5	87	0 / 93	0/0 6	140	9	85-89
79/5–84/ 5	82	0/ 87	0/0 8	131	12	80-84
74/5–79 /5	77	0/ 79	0/0 13	119	19	75-79
69/5-74/5	72	0/ 66	0/ 16	100	25	70-74
64/5-69/5	67	0 / 50	0/ 11	75	17	65-69
59/5-64/5	62	0/ 39	0/ 0 7	58	10	60-64
54/5-59/5	57	0/32	0/0 8	48	12	55-59
49/5-54/5	52	0/24	0/07	36	11	50-54
44/5-49/5	47	0/17	0/05	25	8	45-49
39/5-44/5	42	0/12	0/05	17	7	40-44
34/5-39/5	37	0/07	0/04	10	6	35-39
29/5-34/5	32	0/03	0/03	4	4	30-34

توجه کنید که برای به دست آوردن فراوانی نسبی، فراوانی تجمعی، فراوانی مطلق و فراوانی تجمعی نسبی طبق روال قبلی که توضیح داده‌ایم عمل می‌کنیم .

یادآوری می‌کنیم که ستون فراوانی نسبی (p) نشان دهنده‌ی این است که داده‌ها چه درصدی از کل داده‌ها را تشکیل می‌دهند.

ستون فراوانی تجمعی (CP)، جایگاه افراد در کلاس مورد نظر را نشان می‌دهد. به‌عنوان مثال، کسی که در دسته 89 ـ 85 قرار می‌گیرد از 93 درصد افراد نمره‌ای بیشتر و یا برابر دارد و یا 93 درصد افراد نمره‌ای کمتر یا مساوی با او گرفته‌اند.

12. 4 نمودارهای آماری

روش‌های مرسوم نمایش داده‌ها به صورت هندسی عبارتند از:

الف) نمودار میله‌ای ب) هیستوگرام ج) نمودار چند ضلعی د) نمودار فراوانی تراکمی(اجایو) هـ) نمودار دایره‌ای یا کیکی

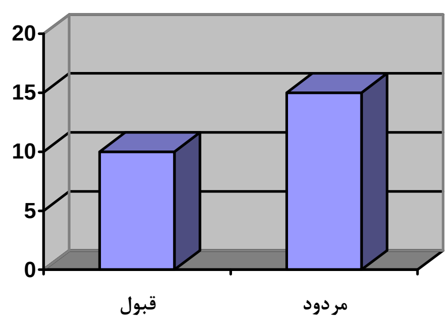
1. 12. 4 نمودار میله ای

این نمودارهنگامی سودمند است که داده‌های جمع‌آوری شده متعلق به متغیرهای گسسته (نظیر قبول و مردود شدن) بوده و با استفاده از مقیاس اسمی اندازه‌گیری می‌شوند. در این نمودار برای نشان دادن طبقات گسسته از مستطیل استفاده می‌شود. ارتفاع میله برابر با فراوانی طبقات است.
مثال: جدول زیر توزیع فراوانی 25 داوطلب که به‌صورت قبول و رد طبقه‌بندی شده‌اند را نشان می‌دهد.

طبقات	فراوانی (f)	درصد (p)
قبول	10	% 40
رد	15	% 60

نمودار میله‌ای متناظر با آن را رسم کنید.

حل:



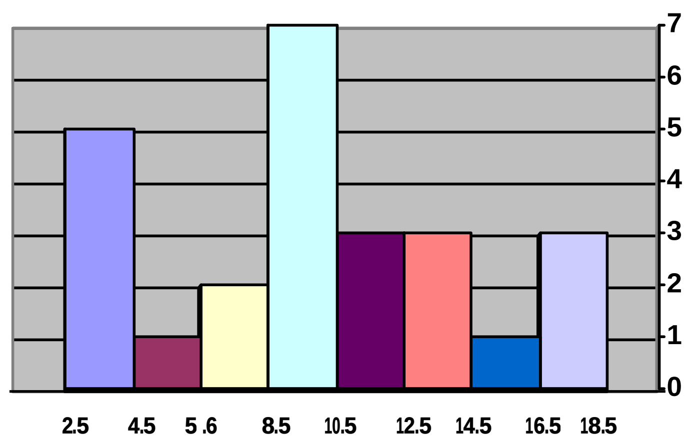
نمودار هیستو گرام

همانند نمودار میله‌ای است و تنها اختلافی که بین این دو نمودار وجود دارد نمایش ستون‌ها است. ستون‌ها در هیستوگرام موجب گردیده تا این نمودار وسیله‌ی مناسبی برای نمایش داده‌های ناشی از اجرای متغیرهای پیوسته و یا متغیرهایی که با استفاده از مقیاس‌های فاصله‌ای و نسبی مورد اندازه‌گیری قرار می‌گیرند باشد. هر ستون نشان دهنده‌ی یک طبقه از اعداد است. حدود واقعی هر طبقه موجب به وجود آمدن معیار پیوسته‌ای بر روی محور افقی (x) می‌شود. عرض هر ستون برابر با طول هر طبقه بوده و ارتفاع هر ستون فراوانی طبقه متناظر را نشان می‌دهد.

جدول. اطلاعات مربوط به 25 دانشجو را در آزمون بارفیکس نشان می‌دهد.

فراوانی تراکمی cf	نقطه میانی xc	فراوانی (f)	کرانه طبقات	حدود طبقات
25	17/ 5	3	16/5-18/5	17-18
22	15/5	1	14/5-16/5	15- 16
21	13/5	3	12/5-14/5	13-14
18	11/5	3	10/5-12/5	11-12
15	9/5	7	8/5-10/5	9-10
8	7/5	2	6/5-8/5	7-8
6	5/5	1	4/5-6/5	5-6
5	3/5	5	2/5-4/5	3-4

نمودار هیستوگرام آن را رسم کنید؟

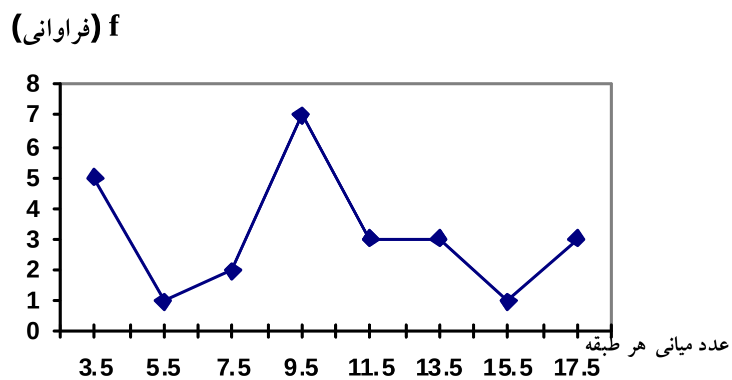


شکل: نمودار هیستوگرام

نمودار چند ضلعی

این نمودارمعمولاً بیشترین کاربرد را دارد؛ و ازتعدادی نقاط که بـه وسـیله خطـوط راسـت بـه هـم متصل می‌شوند تشکیل شده است همچنین این نمودار، برای داده‌های پیوسته و کمی کـاربرد دارد و برای ترسیم آن نمودار، از نقطه وسط هر طبقه استفاده می‌شود. از نقطه وسط هر طبقـه کـه بـر روی محور افقی معین شده است به اندازه فراوانی آن طبقه به موازات محور عمودی بالا می‌رویـم و محل تقاطع را با یک، نقطه مشخص می‌سازیم سپس نقاط حاصل را به وسیله خطوط راست به هم وصل می‌کنیم. این نمودار برای مقایسه یک یا چند گروه از داده‌ها مـورد اسـتفاده قـرار مـی‌گیـرد.

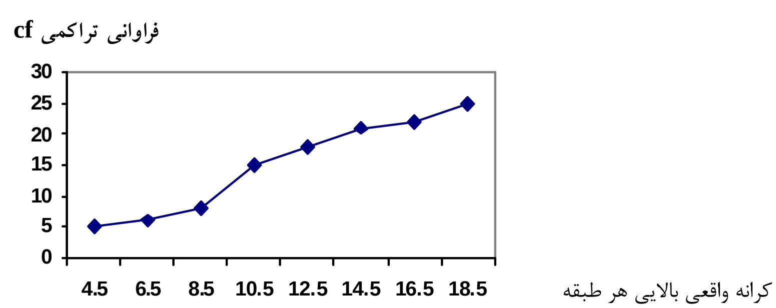
نمودار چند ضلعی متناظر با جدول فراوانی مثال قبل به صورت زیر است:



شکل 3-4: نمودار چند ضلعی

نمودار فراوانی تراکمی یا (اجایو)

برای ترسیم این نمودار محور، عمودی (y) براساس فراوانی تراکمی درجه بندی مـی‌شـود و محـور افقی (X) به جای نقطه میانی بر پایه کرانه‌های واقعی بـالایی هـر طبقـه درجـه‌بنـدی مـی‌شـود. نقطه‌هایی که روی این کرانه‌ها قرار می‌گیرد نشان می‌دهد که چه تعداد و یا چه درصدی از نمره‌ها تا پایان کرانه واقعی بالایی طبقه جای گرفته است. نمودار تراکمی متناظر با جـدول فراوانـی مثـال فوق به‌صورت زیر است:



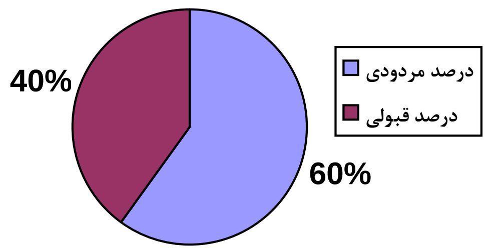
شکل4-4: نمودار فراوانی تراکمی (اجایو)

نمودار دایره ای یا کیکی

این نمودار ساده‌ترین و درعین حال مناسب ترین روش بـرای مقایسـه و نمـایش داده هـای جمـع آوری شده از متغیرهای گسسته و کیفی است. این‌نمودار بر اساس تبدیل سطح دایـره بـه سـطحی معـادل بـا سطح توزیع فراوانی (صد درصد) ترسیم می‌شود. مثلاً درصد داوطلبان قبولی و مردودی را که به ترتیب

40 درصد و 60 درصد هستند محاسبه می‌کنیم. برای این منظور، دایره‌ای ترسیم می‌کنیم که سطح آن برابر با 100 درصد است. باتوجه به اینکه سطح کل دایره 360 درجه است40 پس درصد آن برابر با زاویه‌ای به اندازه144 درجه است. ($144 = 360 \times 40\%$) و 60 درصد آن نیز 216 درجه می‌شود. در حالت کلی برای محاسبه مقدار و درجه یک فراوانی می‌توانیم از فرمول ($d^o = \frac{F}{N}(360^o)$) استفاده نمود.

نمودار دایره‌ای متناظر با اطلاعات فوق به‌صورت زیر است:



شکل: نمودار دایر ه‌ای یا کیکی

معیارهای مرکزی (شاخص های گرایش به مرکز)

معیارهای مرکزی، اندازه تمایل داده‌ها را نسبت به مرکز داده‌ها مشخص می‌کنند که سه نمونه از آن‌ها عبارتنداز: میانگین حسابی ، میانه و مد (نما).

میانگین حسابی

میانگین[1] با حرف $\bar{X}$ نشان داده می‌شود؛ و یکی از قوی‌ترین و مفیدترین اندازه‌ای است که به منظور توصیف گرایش به مرکز، یا معدل توزیع نمره‌های یک گروه از افراد، اشیاء و یا حوادث به کار برده می‌شود. میانگین مناسب‌ترین شاخص تمایل مرکزی برای داده‌های فاصله‌ای و نسبی به شمار می‌رود که از تقسیم جمع کل نمره‌ها بر تعداد آن‌ها به‌دست می‌آید.

الف) در حالتی که داده‌ها طبقه‌بندی نشده باشند از فرمول زیر برای محاسبه میانگین استفاده می‌کنیم:

$$\bar{x} = \frac{\Sigma x_i}{N}$$

$\bar{X} =$ معدل (میانگین)

$N =$ تعداد کل آزمون ها

۱ . Average(mean)

جمع کل آزمون ها $= \Sigma x_i$

مثال نمره‌های 5 دانش‌آموز در آزمون دراز و نشست به شرح زیر ثبت شده است. میانگین حسابی آن‌ها را به‌دست آورید؟

15,20,25,30, 35

حل:

$$\Sigma x_i = 15 + 20 + 25 + 30 + 35$$
$$\overline{x} = \frac{\Sigma x_i}{N} \Rightarrow \overline{x} = \frac{125}{5} \Rightarrow \overline{x} = 25$$

نکته‌ها:

1. حاصل ضرب میانگین در تعداد افراد مساوی است با جمع کل مقادیر مربوط به آن افراد یعنی

$$\overline{x} = \frac{\Sigma x_i}{N} \Rightarrow \Sigma x_i = N\overline{X}$$

2. اگر عدد ثابت c را به کلیه رکوردهای یک گروه اضافه و یا کم نماییم به همان نسبت به میانگین اضافه یا از آن کم می گردد .

به عبارت دیگر، اگر میانگین $X_1, X_2, ...X_N$ برابر $\overline{X}$ باشد آنگاه میانگین $X_1 \pm C, X_2 \pm C, ..., X_N \pm C$ برابر با $\overline{X} \pm C$ است.

3. اگر عدد ثابت c را درکلیه رکوردهای یک گروه ضرب کنیم میانگین نیز در همان عدد ضرب می‌شود.

به عبارت دیگر، اگر میانگین $X_1, X_2, ...X_N$ برابر با $\overline{X}$ باشد آنگاه میانگین $CX_1, CX_2, ..., CX_N$ برابر $C\overline{X}$ است.

4. اگر $\overline{x}$ میانگین داده‌های $x_1, x_2, ..., x_n$ باشد آنگاه میانگین $ax_1 + b, ..., ax_n + b$ به صورت $a\overline{x} + b$ می‌باشد.

مثال: اگر میانگین داده‌های $x_1, x_2, ..., x_n$ برابر 5 باشد میانگین $2x_1 - 2, ..., 2x_n - 2$ را محاسبه کنید؟

حل: چون میانگین $\overline{x} = a\overline{x} - b$ می‌باشد، بنابراین میانگین جدید برابراست با:

$$\overline{x} = (2 \times 5) + (-2) = 10 - 2 = 8$$

5. حاصل جمع جبری انحرافات از میانگین‌ها برابر با صفر می‌باشد.

$$\Sigma(x_i - \overline{x}) = 0$$

6. میانگین داده‌های n ، ... 2 ، 1 برابر با $\frac{n+1}{2}$ است.

7. میانگین داده‌ای $a, a+d ... + a(n-1)d$ که تشکیل یک تصاعد حسابی با قدرنسبت d را می‌دهند برابر است با $\frac{1}{2}[2a + (n-1)d]$

8. اگر میانگین داده‌های $x_1, x_2, ..., x_n$ برابر $\overline{x}$ باشد آنگاه میانگین $x_1+1, x_2+2, ..., x_n+n$ برابر است با $\overline{x} + \frac{n+1}{2}$

مثال: رکوردهای 5 نفر در آزمون بارفیکس به صورت زیر به دست آمده است، نشان دهید که حاصل جمع انحراف از میانگین برابر صفر است؟

12 ، 15 ، 17 ، 17 ، 19

حل:

$$\Sigma(x_i - \overline{x}) = 0$$

$$\overline{x} = \frac{19+17+17+15+12}{5} = 16$$

$$\Sigma(x_i - \overline{x}) = (19-16) + (17-16) + (17-16) + (15-16) + (12-16) =$$

$$(+3) + (+1) + (+1) + (-1) + (-4) = 0$$

نکته: از آنجایی‌که مجموع انحراف از میانگین‌ها صفر است، از حاصل جمع قدر مطلق انحراف از میانگین‌ها استفاده می‌شود؛ که در این صورت باید محاسبات را انجام داد تا نتیجه به‌دست آید.

مثال: جمع جبری قدر مطلق انحرافات از میانگین اعداد زیر را به‌دست آورید؟

x_i = ۲،۳،۷،۸،۱۰

حل:

$$\overline{x} = \frac{10+8+7+3+2}{5} = \overline{x} = \frac{30}{5} \Rightarrow \overline{x} = 6$$

x_i	$x_i - \overline{x}$	$\lvert x_i - \overline{x} \rvert$
10	10-6=4	|10-6|=4
8	8-6=2	|8-6|=2
7	7-6=1	|7-6|=1
3	3-6=-3	|3-6|=3
2	2-6=-4	|2-6|=4
	$\Sigma(x_i - \overline{x}) = 0$	$\Sigma \lvert x_i - \overline{x} \rvert = 14$

مثال: میانگین ده داده آماری برابری 15 محاسبه شده است. پس از محاسبه معلوم شد که باید دو داده 6 و 18 را به آن بیفزائیم، میانگین داده‌های جدید را به‌دست آورید.

حل:

مجموع ده داده آماری $= 15 \times 10 = 150$

مجموع 12داده آماری $= 150 + 6 + 18 = 174$

میانگین 12 داده $= \frac{174}{12} = 14/5$

مثال: میانگین داده‌های آماری 36 ، ... ، 9 ، 6 ، 3 کدام است؟

حل:

$$3+6+9+\ldots+36 = 3(1+2+3+\ldots+12)$$

میانگین داده‌های $1,2,3,\ldots,12$ برابر $\frac{12+1}{2} = 6/5$ است پس میانگین داده‌های مورد نظر $3 \times 6/5 = 19/5$ است.

یا در اینجا می‌توان به صورت مستقیم از رابطه $\frac{3+36}{2} = 19/5$ برای محاسبه میانگین استفاده کرد.

محاسبه میانگین در جدول فراوانی ساده :

برای محاسبه‌ی میانگین در جدول فراوانی ساده باید ابتدا تمام نمرات را در فراوانی مربوط ضرب و سپس مجموع آن‌ها را بر تعداد کل تقسیم کنیم. در این حالت حاصل ضرب فراوانی در نمرات را fx

و مجموع آن را Σfx می‌نامند. که فرمول محاسبهٔ میانگین به‌صورت زیر است.

$$\overline{X} = \frac{\sum fx_i}{N}$$

مثال: اگر جدول زیر در دست باشد میانگین آن به‌صورت زیر حساب می‌شود.

حل:

اعداد (x)	فراوانی (f)
3	1
5	2
6	1
7	3
4	2
	N=9

$$\Sigma fx_i = 3\times1 + 5\times2 + 6\times1 + 7\times3 + 4\times2$$

$$\Sigma fx_i = 48$$

$$\overline{x} = \frac{\Sigma fx_i}{N} = \frac{48}{9} \approx 5/33$$

ب) برای محاسبه میانگین در جدول داده‌های طبقه‌بندی شده باید از نماینده یا عدد میانی هر طبقه استفاده و آن را در فراوانی هر طبقه ضرب، و مجموعه آن‌ها را بر تعداد کل تقسیم کرد؛ به عبارتی دیگر داریم:

$$\overline{x} = \frac{\Sigma fx_c}{N}$$

مثال: میانگین حسابی نمرات زیر را به دست آورید؟

نمرات	18-20	15-17	12-14	9-11
فراوانی	2	2	1	1

حل:

نمرات	فراوانی (f)	نمره میانی (x_c)
18-2	2	19
15-17	2	16
12-14	1	13
9-11	1	10
	N =6	

$$\Sigma fx_c = (2\times19) + (2\times16) + (1\times13) + (1\times10)$$
$$\Sigma fx_c = 93$$
$$\bar{x} = \frac{\Sigma fx_c}{N} = \frac{93}{6} = 15/5$$

ج) برای محاسبه میانگین می‌توان از راه میانگین فرضی نیز عمل کرد که به روش کد گذاری مشهور است. دراین روش یکی از طبقات را که فرض می‌کنیم میانگین در آن واقع شـده اسـت، در ستون جداگانه‌ای که آن را ستون d می نامیم، مساوی صفر قرار مـی‌دهـیم و بـه طبقـه‌ای کـه در ستون d به آن صفر تعلق گرفته است طبقه‌ی مبدأ می‌گوییم، (بهتر است طبقه‌ای را که فراوانی آن از فراوانی دیگر طبقات بیشتر است، به عنوان طبقه مبدأ انتخاب کنیم، البته اگر هم طبقه‌ای دیگـر را به این عنوان مشخص کنیم، در نتیجه محاسبات تغییری ایجاد نمی‌شود).سپس به طبقات بالای طبقه‌ی مبدأ به ترتیب از اولین طبقه بالاتر عدد 1 + به دومین طبقه 2+ و الی آخر عدد مـی‌دهـیم. همچنین در ستونd به طبقات پایین‌تر از طبقۀمبدأ به ترتیب یک طبقه پایین‌تر عدد 1− و دومین طبقه پایین‌تر2- والی آخر عدد می‌دهیم. با این روش، ستون d کامل می‌شود. و ما می‌توانیم ستون دیگری به نام ستون fd تشکیل دهیم و فراوانی هر طبقه را در عدد d همان طبقه ضرب کنـیم ودر مقابل آن در ستون fd می‌نویسیم. با تکمیل ستون fd مجموع این ستون را با در نظر گرفتن اعداد مثبت و منفی محاسبه می‌کنـیم و در فرمـول زیـر بـرای بـه دسـت آوردن میـانگین حقیقـی قـرار می‌دهیم.

$$\bar{x} = x^1 + (\frac{\Sigma fd}{N}) \times i$$

میانگین = $\bar{X}$

عدد میانی طبقه ای که صفر فرضی در ان قرار دارد. = X^1

fd مجموع ستون Σfd =

تعداد کل نفرات = N

به عنوان مثال، جدول زیر را در نظر بگیرید و میانگین آن را از طریق میانگین فرضی به دست آورید.

طبقات	f	d	Fd
90-94	2	+5	10
85-89	3	+4	12
80-84	4	+3	12
75-79	5	+2	10
70-74	6	+1	6
65-69	8	0	0
60-64	7	-1	-7
55-59	4	-2	-8
50-54	3	-3	-9
45-49	5	-4	-20
40-44	2	-5	-10
35-39	1	-6	-6
	n = 50		$\Sigma fd = -10$

$$\bar{x} = x^1 + (\frac{\Sigma fd}{N}) \times i$$

عدد میانی طبقه‌ای که صفر فرضی در آن قرار دارد. $x^1 = \frac{65+69}{2} \Rightarrow x^1 = 67$

$$\bar{x} = 67 + (\frac{-10}{50}) \times 5 = 67 + (-1) \Rightarrow \bar{x} = 66$$

میانه

مقداری که 50 درصد داده‌ها از آن کمتر و 50 درصد داده‌ها از آن بیشتر هستند، **میانه**[1] نامیده می‌شود.

روش تعیین میانه

الف. محاسبه میانه در اعداد طبقه بندی نشده

برای پیدا کردن میانه‌ی اعدادی که تعداد آن‌ها کم است، نیازی به طبقه بندی نیست، داده‌ها را از کوچک به بزرگ مرتب می‌کنیم و چنانچه **تعداد آن‌ها فرد باشد**، میانه عبارت خواهد بود از مقداری که در وسط توزیع قرار دارد و فراوانی آن یک است (عدد تکرار نشده باشد) و اگر **تعداد اعداد زوج باشد** برای تعیین عددمیانی دو عدد وسطی را جمع و بر دو تقسیم می‌کنیم.

مثال: میانه **اعداد ۱۵،۱۴،۱۳،۱۲،۱۰،۹،۷ را به دست آورید؟**

حل:

اعداد را از کوچک به بزرگ مرتب می‌کنیم. 15، 14، 13، 12، 10، 9، 7 چون تعداد فرد باشد، بنابراین میانه برابر 12 است.

مثال: میانه اعداد ۱۲،۱۳ ، 11، 9، 10، 7 را به دست آورید؟

حل: ابتدا داده‌ها را از کوچک به بزرگ مرتب می‌کنیم. چون تعداد داده‌ها زوج است، میانگین دو عدد وسط یعنی 10و11 با میانه برابر است. لذا میانه برابر $\frac{10+11}{2}=10/5$ می‌باشد.

اگر نمره‌ای که در وسط توزیع قرار دارد دارای فراوانی بیش از یک باشد، تشخیص فاصله بین اعداد برای محاسبه میانه ضروری است. فاصله هر عدد، به حدود واقعی همان عدد تعریف شده است. می‌دانیم که دامنه هر عدد یا حدود واقعی هر عدد 0/5 واحد پایین‌تر و 0/5 واحد بالاتر از همان عدد است، بنابراین عدد 5 شامل کلیه اعدادی است که بین 5 /4 تا5/5 قرار دارند، یعنی عدد 5 دارای حد پایین 4/5 و حد بالای 5/5 است به توزیع زیر توجه کنید.

3، 4 ، 4 ، 5، 5 ،5، 5، 6 ، 7

در این توزیع تعداد این نمره‌ها فرد است و میانه در محلی که به وسیله علامت پیکان نشان داده شده قرار دارد. به‌عبارت دیگر میانه در گروه نمره‌های 5 واقع شده است، اما کدام عدد 5.

فرض می‌کنیم توزیع 5 ،5 ،5 ، 5 در طبقه 5 /4 تا5/5 یکسان است بنابراین فاصله هر یک از

[1] .Median

اعداد 5 نسبت به یکدیگر، همان‌طور که در زیر نشان داده شده $\frac{1}{4}$ (یک تقسم تعداد تکرار عدد) یـا (0/25) است. عدد 0/25 را به حد پایینی عدد 5 یعنی 4/5 اضافه می‌کنیم تا به حد بـالایی یعنـی 5/5 برسیم.

4/5 ،4/75،5 ، 5/ 25، 5/5

حالا تمام اعداد را به صورت زیر مرتب می‌کنیم و میانه آن را به دست می‌آوریم.

3 ،4، 4 ، 4/ 5 ، 4/75 ، 5 ، 5/25 ،5/5 ،6، 7

چون تعداد اعداد آن زوج است پس مجموع دو عدد وسطی را بر دو تقسیم می‌کنـیم تـا عـدد میانه به‌دست آید .

$$\text{میانه} = \frac{4/75+5}{2} = 4/875$$

$$\text{میانه} = 4/875$$

روش دیگر برای محاسبه میانه هنگامی که داده‌ها تکراری است:

گاهی در مجموعه داده‌ها بعد از مشخص کردن حل میانه، میانه روی عدد تکراری می‌افتـد در ایـن حالت میانه از رابطه زیر به‌دست می‌آید :

$$\text{میانه} = \text{حد پایین عدد تکراری} + \left(\frac{\frac{N}{2} - \text{تعداد اعداد قبل از عدد تکراری}}{\text{تعداد اعداد تکراری}} \right) \times 1$$

مثال : در مجموعه داده های 4،2 ،6،10،6،11،6،8،6، 12 میانه عدد 6/25 است (Md=6/25) .

مراحل تعیین میانه در حالت تکراری :

اول، اعداد را مرتب می‌کنیم. 2،4،6،6،6،6،8،10،11،12

دوم، حل پایینی عدد تکراری وسطی برابر 5/0می‌شود.

سوم، از رابطه بالا استفاده می‌کنیم و میانه را محاسبه کرده و می‌نویسیم.

$$Md = 5/5 + \left(\frac{(\frac{10}{2} - 2)}{4}\right) \times 1$$

$$Md = 5/5 + 0/75 = 6/25$$

$$Md = 6/25$$

توجه : در این حالت میانه عددی است که بین حد پایین و حد بالا عدد تکراری قرار می‌گیرد.

محاسبه میانه هنگامی که داده‌ها فراوانی نابرابر دارند:

گاهی در مجموعه داده‌ها بعد از مشخص کردن محل میانه، میانه در بین اعدادی که فراوانی نـابرابر دارند قرار می‌گیرد.

مثال : در مجموعه داده‌های 14و14و16و14و16و12و10و20و17و19 میانه عدد 15/2 اسـت یعنـی (Md=15 /2).

مراحل تعیین میانه در حالت فراوانی نابرابر:

اول، اعداد را مرتب می‌کنیم. 20و19و17و16و16و14و14و14و12و10

دوم، محل میانه را تعیین می‌کنیم. $\frac{10+1}{2} = 5/5 =$ محل میانه

سوم، در این مثال میانه بین عدد پنجم و ششم قرار دارد. چون فراوانی این دو عدد (یعنی 14و16) برابر نیستند لذا برای محاسبه میانه باید اختلاف این دو عدد، یعنی عدد 2 (2=14-16) را به نسبت 3 (یعنی فراوانی عدد 14) و به نسبت 2 (یعنی فراوانی عدد 16) تقسیم کنیم.

عدد 14 سه بار تکرار می‌شود. $\frac{2\times 3}{5} = 1/2$

عدد 16 دو بار تکرار می‌شود. $\frac{2\times 2}{5} = 0/8$

بعد می‌توانیم میانه را با دو روش محاسبه کرده و بنویسیم.

Md = عدد کوچک $+1/2 \Rightarrow 14 + 1/2 = 15/2$

Md = عدد بزرگ $-0/8 \Rightarrow 16 - 0/8 = 15/2$

Md = 15/2

ب. میانه در اعداد طبقه بندی شده

چنانچه داده‌ها را به صورت توزیع فراوانی طبقه بندی کرده باشیم میانه‌ی آنها بـه ترتیـب بـا استفاده از الگوریتم زیر به دست می‌آید:

1. ابتدا $\frac{N}{2}$ را محاسبه می‌کنیم. (N تعداد کل فراوانی یا افراد)
2. در ستون فراوانی تجمعی (تراکمی) جدول توزیع فراوانی طبقه بندی شـده، طبقـه‌ای کـه دارای فراوانی تجمعی برابر با $\frac{N}{2}$ یا نزدیکترین عدد بزرگتر از $\frac{N}{2}$ باشد را انتخاب می‌نماییم.
3. حد پایین طبقه مورد نظر را یادداشت می‌نماییم.
4. مقداری را که باید به حد پایینی طبقهٔ میانه اضافه شود، به ترتیب زیر محاسبه می‌نماییم.
5. در نتیجه، مراحل محاسبه میانه را می‌توان با فرمول زیر ارائه نمود:

$$\text{میانه} = \text{حد پایین طبقه میانه} + \left(\frac{\frac{N}{2} - \text{مجموع فراوانی‌های قبل از میانه}}{\text{فراوانی ساده طبقه میانه}} \right) \times \text{فاصله طبقات}$$

$$Mdn = L + \left(\frac{\frac{N}{2} - cf}{f}\right) \times I$$

Mdn = میانه

L = حد پایین طبقه میانه

I = فاصله طبقاتی

Cf = مجموع فراوانی های قبل از طبقه میانه

F = فراوانی ساده طبقه میانه

مثال: میانه جدول زیر را به دست آورید.

x	فراوانی مطلق (f)	فراوانی تجمعی (cf)
35 -39	4	20
30-34	5	16
25 -29	3	11
20-24	4	8
15 -19	2	4
10-14	2	2
	N =20	

حل:

ابتدا $\frac{N}{2}$ را به دست می‌آوریم.

$$\frac{N}{2} = \frac{20}{2} = 10$$

با توجه به عدد به‌دست آمده از پایین، فراوانی را جمع می‌کنیم تا به عدد مورد نظر برسیم آن وقت می‌فهمیم که میانه در کجا قرار دارد. میانه در طبقه 25 – 29 قرار دارد. حالا با توجه به فرمول زیر میانه را به‌دست می‌آوریم.

$$mdn = L + (\frac{\frac{N}{2} - cf}{f}) \times I$$

$$mdn = 24/5 + (\frac{10-8}{3}) \times 5 = 27/8$$

$$mdn = 27/8$$

میانه را وقتی محاسبه می‌کنیم که :

الف. وقت کافی برای محاسبه میانگین نداشته باشیم.

ب. توزیع دارای چولگی قابل ملاحظه باشد؛ یعنی وقتی که یک یا چند نمره در حد افراط و تفریط قرار گیرد.

ج. بخواهیم بدانیم که نتایج اندازه‌گیری در نیمه بالایی توزیع قرار می‌گیرد یا در پایین آن.

د. اندازه‌گیری مقادیر انتهایی توزیع دقیق نباشد.

نما (مد)

نما[1] یکی دیگر از شاخص‌های گرایش به مرکز است؛ یعنی عددی که در توزیع نمره‌ها بیشترین فراوانی را دارد. گاهی اوقات نمرات دارای بیش از یک نماست. نما با داده‌های اسمی به کار برده می‌شود در برآورد سریع مقادیر متوسط (گرایش‌های مرکزی) از نما استفاده می‌شود.

مثال: برای داده‌های ۲،۳،۳،۴،۴،۶،۷ نما برابر با 3و4 می‌باشد. زیرا ماگزیمم فراوانی، مربوط به این دو عدد است. همان‌طوری‌که مثال فوق نشان می‌دهد نما منحصر به فرد نیست.

نکته : بین میانه و میانگین و نما رابطه زیر بر قراراست:

(میانگین) 2-(میانه) (3) = نما   یا $Mo = 3Md - 2\bar{X}$

مثال: اگر میانه دسته‌ای از اعداد 15 و میانگین 10 باشد، نمای آن چقدر است؟

حل:

نما = 3 × (15) – 2 × (10) = 45 – 20 ⇒ نما = 25

برای محاسبه نما در داده‌های دسته‌بندی شده می‌توان از فرمول $MO = La + \frac{1}{2}(I)$ استفاده کرد. در آن Mo = مد ، La = حد پایینی واقعی طبقه با بیشترین فراوانی و I = فاصله طبقات، است.

همچنین می‌توان از فرمول $Mo = L + \left(\frac{d_1}{d_1 + d_2}\right) \times I$ محاسبه نما استفاده کرد که در

[1]. Mode

این فرمول نیز داریم:

مد = MO

L = کرانه پایینی طبقه‌ای که نما در آن قرار دارد. (طبقه‌ای که بیشترین فراوانی دارد)

d_1 = تفاضل فراوانی مطلق طبقه نما دار از طبقه ما قبل.

d_2 = تفاضل فراوانی مطلق طبقه نما دار از طبقه ما بعد.

I = فاصله طبقات.

تذکر: طبقه‌ای که متناظر با فراوانی ماگزیمم است به طبقه نما دار معروف است.

برای برآورد تقریبی نما، در برخی از اوقات مرکز طبقه‌ای را که فراوانی آن حداکثر است در نظر می‌گیریم؛ ولی اشتباه چنین برآوردی معمولاً زیاد است.

مثال ــ نما یا مد جدول زیر را محاسبه کنید.

طبقات	F
18-20	9
15-17	9
12-14	17
9-11	11
6-8	5
3-5	3

حل: طبقه‌ای که بیشترین فراوانی را دارد انتخاب می‌کنیم. در این جدول، طبقه 12-14 دارای بیشترین فراوانی می‌باشد، طبق فرمول بالا حل داریم.

$d_1= 17- 11=6$

$d_2=17 -9 =8$

$$\mathrm{Mo} = \mathrm{L} + \left(\frac{d_1}{d_1 + d_2}\right) \times \mathrm{I}$$

$$\mathrm{Mo} = 11/5 + (\frac{6}{6+8}) \times 3 \Rightarrow \mathrm{Mo} = 12/78$$

روش دیگر : برای پیدا کردن مد در جدول فراوانی دسته‌هایی که فراوانی آن از همه بیشتر است را پیدا می‌کنیم. نشان آن دسته (مرکز آن دسته) مد می‌باشد.

حل مثال بالا : دسته 12-14 بیشترین فراوانی را دارند نشان آن را محاسبه می‌کنیم:

پس مد 13 می‌باشد. $\frac{12+14}{2}=13$

نما را وقتی محاسبه می‌کنیم که :

الف. برآورد سریعی از اندازه‌های مرکزی توزیع لازم باشد.

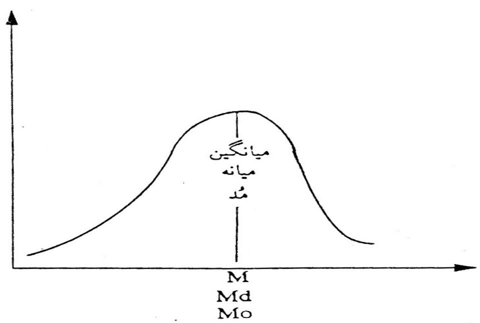
ب. برآورد تقریبی از اندازه‌های مرکزی توزیع لازم باشد.

ج. بخواهیم بدانیم کدام نمره بیشتر تکرار شده است.

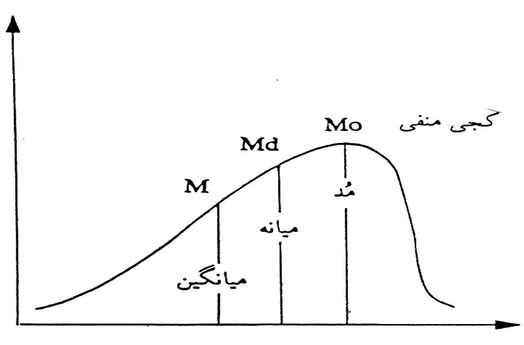
د. مقیاس ما اسمی باشد.

14. 4 میانگین ،میانه و نما در نمودارهای مختلف

الف) اگر میانگین، میانه و نما (مد) برابر باشد (نمودار طبیعی و متقارن یا توزیـع نرمـال) (چـولگی صفر).

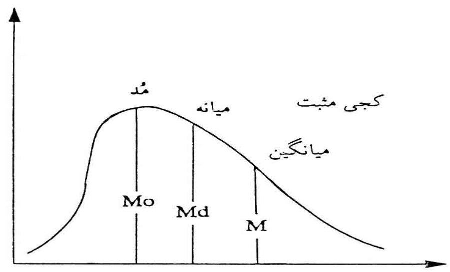


ب) اگر میانگین از میانه و مد کوچکتر باشد (چـولگی منفـی یـا بـه چـپ). در کجـی منفـی میانگین به نمره‌های کوچکتر گرایش دارد. اگر سؤالات یک درس آسان باشد معمولاً منحنی کجـی منفی پیدا می‌کند.



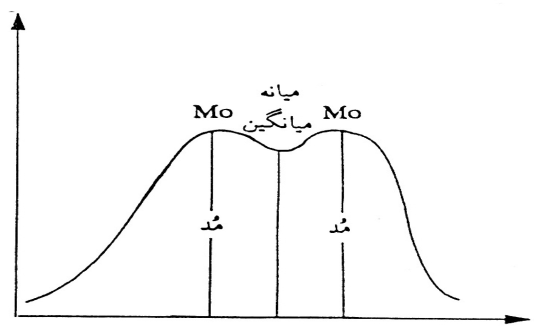
مُد(نما) > میانه > میانگین

ج) اگر میانگین از میانه و مد بزرگتر باشد (چولگی مثبت). در کجی مثبت میانگین به نمرات بزرگتر گرایش پیدا می‌کند. اگر سؤالات یک درس سخت باشد معمولاً منحنی کجی مثبت پیدا می‌کند.



مُد(نما) < میانه < میانگین

د) در صورتی که دو نما یا مد، میانگین و میانه برابر باشد.



معیارهای تغییر پذیری یا پراکندگی

پراکندگی به حدودی که افراد یا نمره‌ها از یکدیگر اختلاف دارند گفته می‌شود. به بیان دیگر تغییرپذیری نمره‌ها عبارت است از میزان اختلاف آن‌ها با میانگین و با یکدیگر. برای مثال فرض کنید دو دسته نمره خام (الف و ب) به شرح زیر وجود دارد:

الف)۵،۷، 9 ب) 3 ، 7 ، 11

هر چند که هر دسته نمره دارای میانگین مساوی 7 است، تغییر پذیری(پراکندگی) در نمره‌های دسته (ب) بیشتراز دسته (الف) است.

پراکندگی را می‌توان از طریق محاسبه دامنه تغییرات، انحراف متوسط، انحراف چارکی، واریانس، انحراف استاندارد (انحراف معیار) و ضریب تغییرات به‌دست آورد.

دامنه تغییرات

که قبلاً گفته شد دامنه تغییرات فاصله بین کوچکترین و بزرگترین نمره در یک توزیع را نشان می‌دهد. لذا چندان قابل اعتماد نبوده و اعتبار کمتری به‌عنوان شاخص پراکندگی دارد و از فرمول زیر به‌دست می‌آید.

$R = (H - L) + \alpha$ یا $R = (X_{max} - X_{min}) + \alpha$

موارد استفاده دامنه تغییرات :

الف. به فوریت شاخصی از اندازه‌های پراکندگی لازم باشد.

ب. اطلاعاتی درباره نمره‌ها در بالا و پایین توزیع نیاز باشد.

مثال: دامنه تغییرات اعداد 13 ، 11 ، 10 ، 9 ، 5 ،3 را به‌دست آورید.

حل:

$R = (13 - 3) + 1 \Rightarrow R = 11$

تذکر: عدد یک در بالا، حاصل نیم نمره حدود واقعی بالایی و پایینی نمرات است.

نکته: هنگامی از دامنه‌ی تغییرات استفاده می‌شود که نمره‌ها دارای پراکندگی خیلی زیادی هستند.

نکته: اگر عدد ثابتی به داده هایی اضافه یا کم شود دامنه تغییرات هیچ تغییری نمی‌کند.

نکته: اگر عدد ثابتی در همه داده ها ضرب یا تقسیم شود دامنه تغییرات نیز در آن عدد ثابت ضرب یا تقسیم می‌شود.

انحراف متوسط (انحراف میانگین)

انحراف متوسط[1] یا انحراف از میانگین به ما کمک می‌کند تا پراکندگی داده‌ها را توصیف کنیم. بدین‌معنا که انحراف از میانگین، شاخصی را برای تعیین مقدار پراکندگی نمره ها از میانگین فراهم می‌کند. انحراف به زبان ساده عبارت است از فاصله هر نمره از میانگین توزیع آن و چون نمره‌ها در اطراف میانگین به صورت مثبت و منفی پراکنده هستند، میزان متوسط واقعی برابر با صفر خواهد بود. متوسط واقعی انحرافات چیزی در مورد توزیع به ما نخواهد گفت بنابراین برای محاسبه متوسط انحرافات از قدر مطلق انحراف از میانگین استفاده می‌شود. اگر انحراف متوسط را با AD نشان دهیم داریم:

$$AD = \frac{\Sigma |x_i - \bar{x}|}{N}$$ طبقه‌بندی نشده

[1] . Average deviation

$$AD = \frac{\Sigma F|x_c - \bar{x}|}{N}$$ دسته‌بندی شده

موارد استفاده انحراف متوسط :

1. تعداد زیادی از نمرات خیلی بزرگ و یا خیلی کوچک موجود که وقتی انحراف آن‌ها از میانگین مجذور می‌شود اثر زیادی بر مقدار انحراف استاندارد بگذارد.
2. شاخص از پراکندگی نسبتاً بدون زحمت لازم باشد.
3. توزیع تقریباً نرمال باشد.

مثال: انحراف متوسط اعداد زیر را به‌دست آورید.

10 ، 8 ، 7 ، 3 ، 2

حل:

$$\bar{x} = \frac{\Sigma x_i}{N} = \frac{30}{5} \Rightarrow \bar{x} = 6$$

$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$
10 -6 =4	$\lvert 10-6 \rvert = 4$
8 -6 =2	$\lvert 8-6 \rvert = 2$
7-6 =1	$\lvert 7-6 \rvert = 1$
3-6 = -3	$\lvert 3-6 \rvert = 3$
2-6= -4	$\lvert 2-6 \rvert = 4$
	$\Sigma \lvert x_i - \bar{x} \rvert = 14$

$$AD = \frac{\Sigma|x_i - \bar{x}|}{N}$$

$$AD = \frac{14}{5} \Rightarrow AD = 2/8$$

مثال:انحراف متوسط نمره های زیر را محاسبه کنید؟

x	f
18- 20	5
10-17	9
12- 14	17
9-11	11
6- 8	5
3- 5	3

حل :

$$\bar{x}=\frac{\sum f_{x_c}}{N}\Rightarrow \bar{x}=\frac{(5\times19)+(9\times16)+(17\times13)+(12\times10)+(5\times7)+(3\times4)}{50}\approx 12/34$$

x	f	X_c عددمیانی	$\lvert x_c-\bar{x}\rvert$	$f\lvert x_c-\bar{x}\rvert$
18-20	5	19	19- 12/34= 6/66	5× 6/66= 33/30
10-17	9	16	16- 12/34= 3/66	9× 3/66 32/94
12-14	17	13	13 -12/34= 0/66	17× 0/66= 11/22
9-11	11	10	10- 12/34= 2/34	11× 2/34= 25/74
6-8	5	7	7- 12/34 = 5/34	5× 5/34 = 26/70
3-5	3 N= 50	4	4- 12/34 = 8/34	3× 8/34= 25/02 $\sum f\lvert x_c-\bar{x}\rvert=154/92$

$$AD=\frac{\sum f\lvert x_c-\bar{x}\rvert}{N}\Rightarrow AD=\frac{154/92}{50}=3/099$$

انحراف چارکی

انحراف چارکی[1]، میزان پراکندگی را در اطراف مرکز توزیع نمره ها نشان می‌دهد. جامعه آماری به چهار قسمت تقسیم می‌شود که چارک اول تا سوم به صورت زیر مشخص می‌شوند:

چارک اول (Q_1): عددی است که از 25 درصد ($\frac{1}{4}$) داده ها بزرگتر و از 75 درصد ($\frac{3}{4}$) داده‌ها کوچکتر است.

چارک دوم (میانه): چارک دوم که با Q_2 نشان داده می شود همان میانه است که از 50 درصد (نصف) داده بزرگتر و از 50 درصد (نصف) داده‌ها کوچکتر است.

چارک سوم (Q_3): عددی است که از 75 درصد ($\frac{3}{4}$) داده‌ها بزرگتر و از 25 درصد ($\frac{1}{4}$) داده‌ها کوچکتر است. از دامنه‌ی تغییرات با ثبات‌تر است و از فرمول زیر محاسبه می‌شود:

---
۱. Quartile Deviation

$$\bar{Q}=\frac{Q_3-Q_1}{2}=\frac{p_{75}-p_{25}}{2}$$

از انحراف چارکی برای معرفی شاخص متوسط از میانه یا تعیین حدود نمره‌های 50 درصدی وسط مورد استفاده قرار می‌گیرد.

یکی از موارد مهم استفاده انحراف چارکی معرفی چولگی (کجی) توزیع است. در توزیع‌های متقارن، فاصله چارک‌های یکم و سوم از میانه یکسان است، اما اگر چولگی توزیع، مثبت (یعنی به راست) باشد مقدار $(Q_3 - Q_2)$ بزرگتر از $(Q_2 - Q_1)$ است و به عکس اگر چولگی توزیع منفی (به چپ) باشد عکس این مطلب صادق خواهد بود، بنابراین داریم:

وقتی چولگی صفر و توزیع متقارن باشد. $Q_3 - Q_2 = Q_2 - Q_1$

وقتی چولگی مثبت باشد. $Q_3 - Q_2 > Q_2 - Q_1$

وقتی چولگی منفی باشد. $Q_3 - Q_2 < Q_2 - Q_1$

نکته: منحنی توزیع نمرات یک کلاس با گرایش مثبت(چولگی مثبت) معرف یک کلاس ضعیف و کلاس با گرایش منفی (چولگی منفی) معرف یک کلاس با سطح قوی‌تر است.

موارد استفاده از انحرف چارکی:

1. برای معرفی شاخص متوسط، از میانه استفاده شده باشد.
2. توزیع نرمال نباشد.
3. چند نمره خیلی بزرگ و یا خیلی کوچک داشته باشیم یا چولگی توزیع شدید باشد.
4. حدود نمره‌های 50 درصد وسط لازم باشد.

الف) محاسبه انحراف چارکی در اعداد طبقه بندی نشده

برای محاسبه انحراف چارکی ابتدا داده‌ها را از کوچکتر به بزرگ مرتب می‌کنیم، سپس میانه توزیع را به‌دست می‌آوریم (Q_2). بعد از این کار تعدادی از اعداد در سمت راست میانه و تعدادی در سمت چپ آن واقع می‌شوند. با بدست آوردن میانه اعداد سمت راست چارک سوم (Q_3) و با بدست آوردن میانه اعداد سمت چپ چارک اول (Q_1) بدست می‌آید. مطابق فرمول، تفاضل چارک اول و سوم را نصف می‌کنیم تا انحراف چارکی به دست آید.

$$\bar{Q}=\frac{Q_3-Q_1}{2}$$

مثال: انحراف چارکی توزیع 8 ، 12، 4 ، 9 ، 5، 11، 10، 7 ،6 کدام است؟

حل: اعداد را از کوچک به بزرگ مرتب می‌کنیم و آنگاه میانه آن را به دست می‌آوریم (Q_2). در این توزیع $Q_2 = 8$ می‌باشد. و اعداد سمت راست میانه ، عدد وسط آن (Q_3) را بدست می‌آوریم که برابر 10/5 می‌باشد. از اعداد سمت چپ، عدد وسطی آن (Q_1) را به‌دست می‌آوریم که برابر 5/5

می‌باشد.

4 ، 5 ،6 ،7 ، 8، 9 ، 10 ،11 ، 12

9.10.11.12

Q_3=10/5 $\quad\quad \overline{Q}=\frac{Q_3-Q_1}{2}=\frac{10/5-5/5}{2}=2/5$

4 ،۶،۵ ، 7

Q_1=5/5

ب. محاسبه انحراف چارکی در اعداد طبقه بندی شده

برای محاسبه چارک اول و سوم ابتدا فراوانی تراکمی داده را در جدول به دست آورده. سپس از روابط زیر چارک‌ها را به دست می‌آوریم.

$$Q_3=L+\left(\frac{\frac{3N}{4}-cf_i}{f_i}\right)\times I \text{ و } Q_1=L+\left(\frac{\frac{N}{4}-cf_i}{f_i}\right)\times I$$

Q_1 = چارک اول

Q_3 = چارک سوم

Cf_i= فراوانی تراکمی طبقه ما قبل

L= کرانه پایینی که Q_1 و Q_3 در آن قرار دارد.

I= فاصله طبقات

مثال: اطلاعات زیر در مورد نمرات درس حرکت شنا سی 40 دانشجوی تربیت بدنی است انحراف چارکی (چارک متوسط) آن را بدست آورید؟

حل:

طبقات	F	CF_i
18-20	10	40
15-17	5	30
12-14	15	25
9-11	5	10
6-8	5	5
	N =40	

1ـ محاسبه چارک اول (Q_1) یا نقطهٔ درصدی 25:

$$P_{25}=\frac{25}{100}N=\frac{N}{4}\Rightarrow P_{25}=\frac{40}{4}=10$$

نگاه می‌کنیم که نمره 10 در کدام طبقه قرار می‌گیرد. (طبقه 11-9) طبق فرمول بالا حل می‌کنیم.

$$Q_1=L+(\frac{\frac{N}{4}-cf_i}{f_i})\times I\Rightarrow Q_1=8/5+(\frac{10-5}{5})\times 3\Rightarrow Q_1=11/5$$

2- محاسبه چارک سوم (Q_3) یا نقطه درصدی 75 :

$$P_{75}=\frac{75}{100}N=\frac{3}{4}N\Rightarrow P_{75}=\frac{3}{4}\times 40=\frac{120}{4}=30$$

طبقه مورد نظر 17-15 می‌باشد چون 30 در آنجا قرار می‌گیرد.

$$Q_3=L+(\frac{\frac{3N}{4}-cf_i}{f_i})\times I\Rightarrow Q_3=14/5+(\frac{30-25}{5})\times 3\Rightarrow Q_3=17/5$$

بنابراین انحراف چارکی برابر است با:

$$\bar{Q}=\frac{Q_3-Q_1}{2}\Rightarrow \bar{Q}=\frac{17/5-11/5}{2}\Rightarrow \bar{Q}=3$$

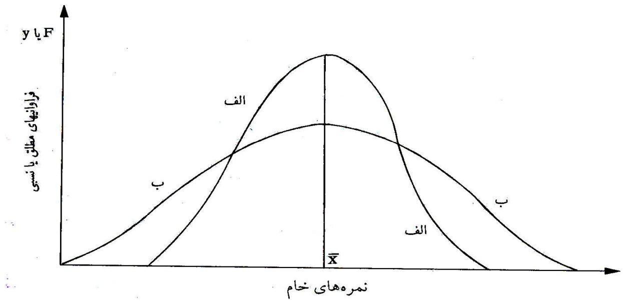
واریانس

یکی دیگر از اندازه‌های پراکندگی که تمام نمرات روی آن اثر می‌گذارند، واریانس[1] است. واریانس مقیاسی است که نشان می‌دهد که داده‌ها حول میانگین چگونه پخش شده‌اند. اگر یکی از نمره‌ها تغییر کند واریانس نیز تغییر می‌کند. به این علت از اهمیت بیشتری برخوردار است. واریانس در واقع میانگین مجذور انحرافات می‌باشد. واریانس کمتر بدین معنا است که انتظار می‌رود اگر نمونه‌ای از توزیع مزبور انتخاب شود مقدار آن به میانگین نزدیک باشد.

شکل 6-4 دو کلاس الف و ب را نشان می‌دهد که از نظر گرایش مرکزی مشابه ولی از جهت تغییر-پذیری نمره‌ها متفاوت هستند. همان طوری که در شکل مشاهده می‌کنید نمر ه‌های کلاس الف بیشتر در حدود گرایش مرکزی خود که با یک خط عمود نشان داده شده است متمرکز هستند. در حالی در کلاس ب نمره‌های بیشتری در فاصله های دورتر از شاخص گرایش مرکزی خود واقع شده‌اند.کلاس الف را متجانس‌تر می‌گویند.

---

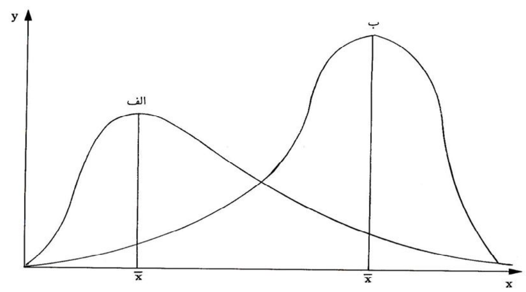
[1].variance



شکل 6-4: تفاوت واریانس های در دو گروه با میانگین های مساوی

شکل 7-4 دو منحنی کج را نشان می‌دهد. هر دوی این منحنی‌ها فاقد ویژگی قرینگی نسبت به نقطه مرکزی خود هستند. در منحنی الف، گرایش منحنی به سمت راست یا نمره‌های بالاست. این گرایش مثبت یا خوب می‌نامند. منحنی ب را دارای گرایش بد، به چپ یا منفی گویند.

منحنی توزیع نمرات یک کلاس با گرایش مثبت معرف یک کلاس ضعیف و کلاس با گرایش منفی معرف یک کلاس با سطح قوی‌تر است.



شکل 7-4: توزیع در دو گروه با میانگین و واریانس های متفاوت

فرمول واریانس به صورت حاصل تقسیم مجموع مجذور انحرافات به تعداد کل اعداد می‌باشد.

برای محاسبه واریانس اعداد طبقه‌بندی نشده از فرمول زیر استفاده می شود:

$$S^2 = V = \frac{\Sigma(x_i - \bar{x})^2}{N} \quad \text{یا} \quad S^2 = V = \frac{\sum X_i^2}{N} - \left(\frac{\sum X_i}{N}\right)^2 \quad \text{و} \quad S^2 = V = \frac{\Sigma x_i^2}{N} - \bar{x}^2$$

واریانس داده‌های طبقه‌بندی شده از فرمول زیر به‌دست می‌آید:

$$S^2 = V = \frac{\Sigma f\,(x_C - \bar{x})^2}{N}$$

تذکر: در نمونه‌های کوچک به جای N از N-1 در مخرج کسر استفاده می‌شود.

نکته: واریانس زمانی صفر است که تمام نمرات با هم برابر باشند.

نکته: هر چه واریانس به صفر نزدیکتر باشد پراکندگی بین داده‌ها کمتر است.

مثال: واریانس داده‌ای آماری۱۵،۱۳،۱۲ ،۱۲،۹،۸،۷،۴ را بیابید.

X_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
4	-6	36
7	-3	9
8	-2	4
9	-1	1
12	+2	4
12	+2	4
13	+3	9
15	+5	25
$\Sigma x_i = 80$	$\Sigma(x_i - \overline{x}) = 0$	$\Sigma(x_i - \overline{x})^2 = 92$

حل:

$$\overline{x} = \frac{\Sigma x_i}{N} = \frac{80}{8} \Rightarrow \overline{x} = 10$$

$$S^2 = v = \frac{\Sigma(x_i - \overline{x})^2}{N} = \frac{92}{8} \Rightarrow S^2 = 11/5$$

مثال : واریانس توزیع زیر را محاسبه کنید؟

x	f
18-20	4
15-17	2
12-14	1
9-11	1
6-8	2
3-5	3

حل:

$$\bar{x}=\frac{\sum f_{x_c}}{N}=\frac{(4\times19)+(2\times16)+(1\times13)+(1\times10)+(2\times7)+(3\times4)}{13}$$

$$\bar{x}=\frac{157}{13}\approx\ 12$$

x	f	$(x_c-\bar{x})$	$x_c-\bar{x}$	$(x_c-\bar{x})^2$	$f_i(x_c-\bar{x})^2$
18-20	4	19	19-12=7	$7^2=49$	4× 49=196
15-17	2	16	16- 12=4	$4^2=16$	2×16 = 32
12-14	1	13	13-12=1	$1^2=1$	1×1=1
9-11	1	10	10-12= -2	$(-2)^2=4$	1× 4=4
6-8	2	7	7-12= -5	$(-5)^2=25$	2×25=50
3-5	3 N=13	4	8-12= -8	$(-8)^2=64$	3× 64=192 $f(x_c-\bar{x})^2=475$

$$s^2=\frac{\sum f(x_c-\bar{x})^2}{N}=\frac{475}{13}=36/5$$

انحراف استاندارد (انحراف معیار)

انحراف استاندارد[1] شاخص آماری اصلی پراکندگی نمره‌ها نسبت به میانگین می‌باشد که نه تنها پراکندگی نمرات را به‌طور دقیق اندازه‌گیری می‌کند، بلکه کمیتی است که در محاسبه و تفسیر بسیاری از موارد از آن استفاده می‌شود.

موارد استفاده از انحراف استاندارد(انحراف معیار) :

1. وقتی معتبرترین اندازه‌گیری پراکندگی لازم باشد.
2. محاسبات بعدی آماری مانند ضریب همبستگی لازم باشد.
3. تعبیر و تفسیر نمره‌ها با در نظر گرفتن خم نرمال لازم باشد.

برای محاسبه انحراف استاندارد، واریانس را محاسبه نموده و جذر آن را محاسبه می‌کنیم. اگر

---

۱ . Standard Deviation

انحراف استاندارد را با S نشان دهیم آنگاه داریم:

$$S=\sqrt{V}=\sqrt{\frac{\sum (X_i-\overline{X})^2}{N}} \quad \text{یا} \quad S=\sqrt{V}=\sqrt{\frac{\sum x_i^2}{N}-\left(\frac{\sum x_i}{N}\right)^2} \quad \text{یا} \quad S=\sqrt{V}=\sqrt{\frac{\sum X_i^2}{N}-\overline{x}^2}$$

$$S=\sqrt{V}=\sqrt{\frac{\sum f(X_c-\overline{X})^2}{N}}$$ برای اعداد طبقه‌بندی شده

لازم به ذکر است که فرمول فوق برای محاسبه‌ی انحراف استاندارد در جمعیت‌های بزرگ به کار می‌رود، در نمونه های کوچک به جای n از n -1 در مخرج کسر استفاده می‌شود. لذا نمونه‌های کوچک بهتر است از فرمول‌های زیر استفاده شود:

$$s=\sqrt{V}=\sqrt{\frac{N\Sigma x_i{}^2-(\Sigma x_i)^2}{N-1}}$$

در حالتی که داده‌ها دارای فراوانی باشند و داده‌ها طبقه بندی شده باشند از فرمول زیر استفاده می‌شود.

$$s=\sqrt{V}=\sqrt{\frac{N\Sigma x_c{}^2-(\Sigma x_c)^2}{N}}$$

دراین حالت اگر داده‌ها طبقه بندی نشده باشد از فرمول زیر استفاده می‌کنیم:

$$s=\sqrt{V}=\sqrt{\frac{N\Sigma f x_i{}^2-(\Sigma x f_i)^2}{N(N-1)}}$$

مثال: اگر در یک جامعه ی آماری $\sum x_i = 15$ و $\sum x_i^2 = 55$ و $n=5$ باشد انحراف معیار چقدر است؟

$$\overline{x}=\frac{\sum x_i}{n}=\frac{15}{5}=3 \Rightarrow S=\sqrt{V}=\sqrt{\frac{\sum x_i^2}{n}-\overline{x}^2}=\sqrt{\frac{55}{5}-(3)^2}=\sqrt{11-9}=\sqrt{2}$$

مثال: انحراف استاندارد داده‌های جدول زیر را بیابید.

x_i	f	fx_i	$f x^2$
42	1	42	1764
41	1	41	1681
40	2	80	3200
39	3	117	4563
38	4	152	5776
36	3	108	3888
35	7	245	8575
34	4	136	4624
33	2	66	2178
32	3	96	3072
31	2	62	1922
30	3	90	2700
29	1	29	841
	N=40	$\Sigma fx_i = 1412$	$\Sigma fx_i^2 = 44784$

حل:

$$s = \sqrt{V} = \sqrt{\frac{N\Sigma fx_i^{\,2} - (\Sigma fx_i)^2}{N(N-1)}} = \sqrt{\frac{40(44784) - (1412)^2}{40(39)}} = \sqrt{\frac{2010400 - 1993744}{40(39)}} = \sqrt{\frac{16545}{1560}} =$$

$\sqrt{10/677} = s = 3/26$ انحراف معیار

نکته: انحراف استاندارد زمانی صفر می‌شود که تمام نمرات با هم برابر باشند .

تفسیر انحراف استاندارد

انحراف استاندارد، شاخص پراکندگی یا تغییرپذیری یک سلسله‌ی اعداد است. می‌توان گفت که هر قدر انحراف استاندارد بزرگتر باشد، تغییر پذیری اعداد نیز بیشتر و هر چه کوچکتر باشد پراکندگی نیز کمتر است. به درستی نمی‌توان گفت که کدام انحراف استاندارد کوچک و کدام بزرگ است. بزرگ یا کوچک بودن انحراف استاندارد امری نسبی است؛ یعنی بزرگتر و کوچکتر در مقام مقایسه با گروه یا دسته‌ی همگن دیگری است که با آن گروه مقایسه می‌شود. مقدار انحراف استاندارد بستگی به واحد اندازه‌گیری دارد برای مثال ممکن است از صفر تا 100 و یا از صفر تا 20 باشد.

بنابراین، حتی در مواقعی که موضوع مورد اندازه‌گیری یکسان، اما مقیاس‌ها مختلف است نمی‌توان بزرگی یا کوچکی انحراف استاندارد را به طور مطلق تعیین کرد.

یکی از محاسن عمده‌ی انحراف استاندارد، رابطه‌ای است که بین واحد انحراف استاندارد و طرز قرارگرفتن نمره‌ها در منحنی طبیعی موجود است. در فصل‌های بعدی درباره آن توضیح می‌دهیم.

اثر تغییر مقیاس بر واریانس و انحراف استاندارد

هرگاه نقطه شروع یک مقیاس اندازه‌گیری به وسیله اضافه کردن یا کسرکردن یک عدد ثابت به هر یک از نمره‌های دیگر در توزیع تغییر پیدا کند، واریانس و انحراف استاندارد تغییر پیدا نمی‌کند. چنانچه واحد اندازه‌گیری را به وسیله ضرب یا تقسیم هر یک از نمره‌های گروه تغییر دهیم، واریانس جدید با ضرب یا تقسیم واریانس قدیم در مجذور آن عدد ثابت بدست می‌آید.

واریانس جدید $s_{x^1}^2 = c^2 s_x^2$

واریانس جدید $s_x^2 = \frac{s_x^2}{c^2}$

چنانچه واحد اندازه‌گیری را در نمره ثابت ضرب یا تقسیم کنیم، انحراف استاندارد جدید برابر خواهد شد با انحراف استاندارد قدیم ضرب یا تقسیم برقدر مطلق همان عدد ثابت.

انحراف استاندارد جدید $s_{x^1} = |c| s_x$

انحراف استاندارد جدید $s_{x^1} = \frac{s_x}{|c|}$

مثال 1 : اگر واریانس درس سنجش و اندازه‌گیری 4 باشد و اگر همه‌ی نمرات در 3 ضرب شود، واریانس جدید را محاسبه کنید؟

$s_{x^1}^2 = c^2 s_x^2 \Rightarrow s_{x^1}^2 = 3^2 \times 4 = 9 \times 4 = 36$

مثال 2 : اگر انحراف استاندارد درس ریاضیات 8 باشد و هر کدام از نمرات را در 2 ضرب کنیم، انحراف استاندارد جدید را محاسبه کنید؟

$s_{x^1} = |c| s_x \Rightarrow s_{x^1} = |2| \times 8 = 16$

مثال 3 : اگر انحراف استاندارد درس یادگیری حرکتی 4 باشد و عدد 3- را در تمام نمرات ضرب کنیم، انحراف استاندارد جدید را محاسبه کنید؟

$S_{x^1} = |c| S_x \Rightarrow S_{x^1} = |-3| \times 4 = 3 \times 4 = 12$

مثال 4 : واریانس درس حرکت‌شناسی 16 می‌باشد، و هر یــک از نمــرات را بــر 2 تقســیم کنیم، واریانس جدید را محاسبه کنید؟

$$s^2_{x^1} = \frac{s_x^2}{c^2} \Rightarrow s^2_{x^1} = \frac{16}{2^2} = \frac{16}{4} = 4$$

مثال 5 : اگر انحراف استاندارد درس آمار 6 باشد، و هر کــدام از نمــرات را بــر 3 تقیســم کنیم، انحراف استاندارد را محاسبه کنید؟

$$s_{x^1} = \frac{s_x}{|c|} \Rightarrow s_{x^1} = \frac{6}{|3|} = 2$$

ضریب تغییرات

ضریب تغییرات[1] شاخصی است برای مقایسه پراکندگی دو ویژگی از یک گـروه بـا یکـدیگر و تنهـا شاخص پراکندگی است که بدون واحد است و فقط برای داده‌های مثبت تعریف مـی‌شـود. مقیـاس نسبی بالاترین مقیاس اندازه‌گیری است. در تربیت‌بدنی کـه مقیـاس نسـبی اسـتفاده زیـادی دارد، پژوهشگران غالباً علاقه‌مندند پراکندگی یک ویژگی در یک نمونه را با پراکندگی ویژگـی در همـان نمونه مقایسه کنند.

برای مثال آیا تغییرات میزان قد یک گروه نوجوان، همانند تغییرات وزن آنهاست، البته نمی‌تـوانیم قد و وزن را به طور مستقیم مقایسه کنیم، اما می‌توانیم ضریب نسبی یک گروه را در رابطه با توزیع قد و وزن مقایسه کنیم، تغییرات (v) گـاهی اوقـات بـه عنـوان ضـریب نسـبی واریـانس شـناخته شده‌است.

در محاسبه ضریب تغییرات (C.V)، از میانگین در انحراف معیار یک توزیع استفاده می‌کنـیم. فرمول محاسبه ضریب تغییرات به شرح زیر است:

$$CV = \frac{SD}{\overline{X}}(100) \quad \text{یا} \quad CV = \frac{100SD}{M}$$

نکته: در ضریب تغییرات واحد اندازه‌گیری حذف می شود و می‌توان بـا ایـن شـاخص پراکنـدگی را برای دو جامعه با مقیاس‌های اندازه‌گیری متفاوت سنجش کرد.

خواص ضریب تغییرات :

1. اگر همه‌ی داده‌ها با هم برابر باشند ضریب تغییرات صفر است.
2. اگر همه‌ی داده‌ها را در یک عدد ثابت (مثبت) ضرب کنیم، ضریب تغییرات تغییر نمی‌کند.

---
[1] .Coefficient of variation

3. اگر به همه‌ی داده ها یک عدد مثبت اضافه کنیم، ضریب تغییرات جدید کوچکتر از ضریب تغییرات داده‌های اولیه است.

4. اگر به همه‌ی عدد را از یک عدد ثابت کم کنیم به مقدار نامعلومی افزایش می‌یابد.

مثال: در یک نمونه از نوجوانان، میانگین قد 150 سانتی متر و انحراف معیار 20 سانتی‌متر است. در همین گروه میانگین وزن 50 کیلوگرم و انحراف معیار وزن 15 کیلوگرم است. در کدام‌یک از متغیرها، ضریب تغییر بزرگ‌تراست؟

حل:

قد $CV = \frac{100\,SD}{M} \Rightarrow CV = \frac{100 \times 20}{150} = \frac{2000}{150} = \%13/3$

وزن $CV = \frac{100 S.D}{M} \Rightarrow CV = \frac{15 \times 100}{50} = \frac{1500}{50} = \%30$

فصل ۲

توزیع نرمال

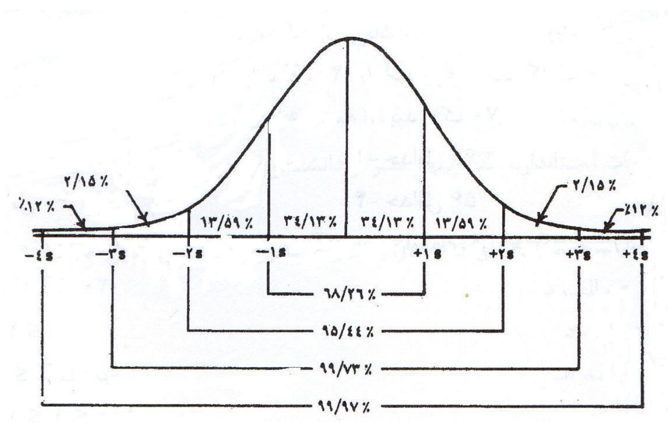
مقدمه

«منحنی طبیعی» مهمترین الگوی علمی در آمار به شـمار مـی‌رود . بیشـتر روش‌هـای محاسـبه و تجزیه و تحلیل آماری بر مبنای خصوصیات ریاضی این منحنی ساخته شـده‌انـد. منحنـی طبیعـی شکلی شبیه به زنگوله دارد و متقارن است.

در این منحنی، میانگین، میانه و نما در بالاترین ارتفاع منحنی برهم منطبقنـد و خطـی کـه از این نقطه موازی با محور yها رسم می‌شود محور تقارن منحنی است. این منحنی بر اساس متغیـر مورد نظر روی محور X و فراوانی‌ها روی محور Y رسم می‌شود.

در وسط منحنی فراوانی‌ها بیشتر است و هر چـه بـه دو طـرف منحنـی نزدیـک مـی‌شـویم، از فراوانی‌ها کاسته می‌شود، ولی هیچگاه به صفر نمی‌رسد. بنابراین، امتداد منحنی با محـور X هرگـز تماس حاصل نمی‌کند. دقیق‌تر آنکه محورx ها مجانب افقی نمودار این منحنی است. یکی از خواص بسیار مهم منحنی طبیعی درصدهای زیر منحنی است که همیشه ثابت هستند.

مساحت سطح زیر منحنی طبیعی برابر در نظر گرفته می‌شـود . بنـابراین، سـطح زیـر منحنـی پائین‌تر از میانگین برابر است با 0/5 و سطح زیـر منحنـی بـالاتر از میـانگین نیـز مسـاوی بـا 0/5 می‌باشد. به عبارت دیگر 50% پائین‌تر از میانگین و 50% بالاتر از میانگین قرار دارند. دامنـه نمـرات Z بین 5 و 5ـ می‌باشد، لیکن از آنجایی کـه 99/72 درصـد افـراد در جامعـه طبیعـی Z در دامنـه نمرات 3 و3ـ قرار دارد، لذا معمولاً دامنه نمرات Z را بین 3 و 3ـ در نظر می‌گیرند.



شکل 1-5. نمودار درصد های زیر منحنی طبیعی استاندارد بر حسب واحد های استاندارد z

درصدهای زیر منحنی اساس بسیاری از تجزیه و تحلیل‌های آماری می‌باشد .

برای مثال، می‌توان جایگاه فردی را در یک امتحان و در گروه خود مشخص نمود. یعنی با استفاده از درصدهای زیر منحنی می‌توانیم بگوییم که این فرد از چند درصد افراد گروه خود بالاتر و یا پایین‌تر است. روش این کار در مبحث نقاط درصدی و تهیه نورم‌های مختلف توضیح داده می‌شود.

نمرات استاندارد

همانطور که می‌دانیم، ارزشیابی بر مبنای مقایسه انجام می‌شود، ولی چنانچه مقادیر به دست آمده از اندازه‌گیری‌ها مبنای مشترک نداشته باشند، امر مقایسه امکان‌پذیر نخواهد بود. پس برای مقایسه نمره‌ها باید آن‌ها را به‌صورت نمره نسبی در آوریم و هر کدام را نسبت به شاخص‌های اساسی گروهی که در آن قرار می‌گیرند، بیان کنیم. بهترین روش تبدیل نمرات خام به نمرات نسبی آن است که هر نمره را با مقیاس میانگین و انحراف استاندارد خود نشان دهیم.

این نمرات به نمرات استاندارد یا Z معروف هستند که از فرمول زیر به‌دست می‌آیند:

$$Z = \frac{X - \overline{X}}{S}$$

در فرمول فوق x نمره خام، $\overline{X}$ میانگین، s انحراف استاندارد (انحراف معیار) و z نمره استاندارد است. برای محتوا بخشیدن به نمرات استاندارد، z آنها را روی محور x در زیر منحنی طبیعی قرار می‌دهیم به این ترتیب که در محل میانگین، عدد صفر می‌گذاریم و هر طرف از منحنی به سه انحراف استاندارد (انحراف معیار) تقسیم می‌شود در حقیقت، نمرات استاندارد z، میانگینی برابر صفر و انحراف استانداردی برابر با یک دارند.

چند نکته مهم:

1. هرگاه یک عدد ثابت به همه‌ی نمرات اضافه شود، به میانگین نیز همان عدد ثابت اضافه می‌شود.
2. هرگاه یک عدد ثابت از همه‌ی نمرات کسر شود، از میانگین نیز همان عدد ثابت کسر خواهد شد.
3. هرگاه یک عدد ثابت در همه‌ی نمرات ضرب شود، میانگین نیز در آن عدد ضرب خواهد شد.
4. هرگاه همه نمرات بر یک عدد ثابت تقسیم شوند، میانگین نیز بر آن عدد ثابت تقسیم خواهد شد.
5. هرگاه یک عددثابت به همه نمرات اضافه شود، درانحراف استاندارد تغییری ایجاد نخواهد شد.
6. هرگاه یک عدد ثابت از همه نمرات کسرشود، در انحراف استاندارد تغییری ایجاد نمی‌شود.
7. هرگاه یک عدد ثابت در همه نمرات ضرب شود، در انحراف استاندارد نیزضرب خواهد شد.
8. هرگاه همه نمرات بریک عدد ثابت تقسیم شوند، انحراف استاندارد نیز بر آن عدد تقسیم می‌شود.

کاربرد نمرات استاندارد Z

برای مقایسه مقادیر مختلف درمقیاس‌های گوناگون باید آن‌ها را به نوعی از مقیاس واحد تبدیل کنیم تا امرمقایسه امکان‌پذیر شود. در این خصوص، از نمرات استاندارد z استفاده می‌کنیم. کاربرد دیگر نمرات استاندارد z، برای تعیین جایگاه افراد در گروه خود می‌باشد که دراین خصوص، از درصدهای زیر منحنی طبیعی استفاده می‌شود. برای روشن‌تر شدن نوع کاربرد نمرات استاندارد z به مثال‌های زیر توجه کنید.

مثال: فردی دریک آزمون پرش طول با میانگین 150 سانتی متر و انحراف استاندارد 20، رکورد 160 سانتی‌متر و در آزمون بارفیکس بامیانگین 12 و انحراف استاندارد 2، رکورد 14 کسب کرده است. مشخص کنید که درکدام آزمون قوی‌تر بوده است.(منحنی طبیعی فرض شود.)

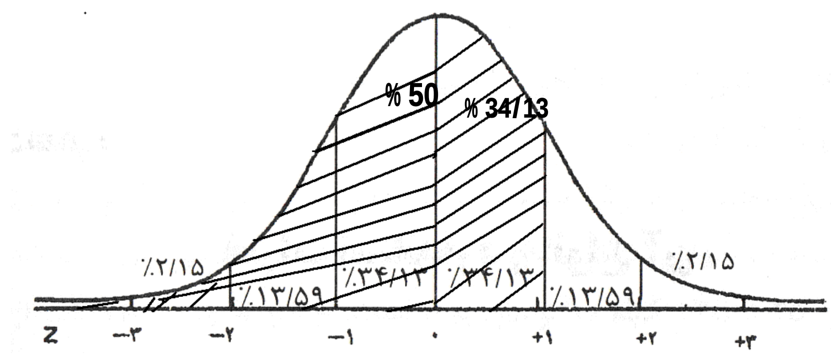
پرش طول $z_1 = \frac{x_1 - \bar{x}_1}{s_1} = \frac{160-150}{20} = \frac{10}{20} = \frac{1}{2}$

کشش بارفیکس $z_2 = \frac{x_2 - \bar{x}_2}{s_2} = \frac{14-12}{2} = \frac{2}{2} = 1$

بنابراین چون z_2 بیشتر z_1 است، مشخص می‌شود که این فرد درآزمون کشش بارفیکس قوی‌تر است.

مثال: فردی درآزمون درازونشست با میانگین 30 وانحراف استاندارد 10، رکورد 40 به دست آورده است. جایگاه وی رادر گروه مشخص کنید. (منحنی رکوردها طبیعی فرض می‌شود).

$$z_1 = \frac{x - \bar{x}}{s} = \frac{40 - 30}{10} = \frac{10}{10} = 1 \Rightarrow z = 1$$



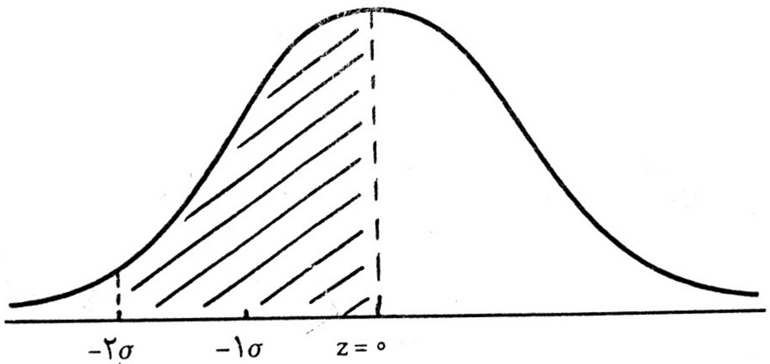
بنابراین، می‌توان گفت که وی یک انحراف استاندارد بالاتر از میانگین قرار دارد و با توجه به درصدهای زیر منحنی طبیعی، وی از 84/13 درصد افراد گروه خود بالاتر است.

50 + 34/13 = 84/13 درصد

مثال: دریک کلاس فوتبال میانگین نمرات 42 وانحراف استاندارد 3 بوده است. چنددرصد افراد نمرات بین 36 و میانگین گرفته‌اند؟

حل: نمرات z نمره خام را به‌دست می‌آوریم .

$$z_1 = \frac{x - \bar{x}}{s} = \frac{36 - 42}{3} = \frac{-6}{3} = -2$$



علامت منفی برای z، نشان می‌دهد که نمره خام مورد نظر پائین‌تـراز میـانگین قـرار داشـته و درسمت چپ آن قرار می‌گیرد. پس ما مساحتی از منحنی که بـین 2 ـــ و میـانگین اسـت را لازم داریم.

حدود 48 درصد افراد نمرات بین میانگین و نمره خام 36 بدست آورده‌اند.

$\%13/59 + \%34/13 = 47/72$

تذکر مهم:

در آزمون‌های تربیت بدنی دونوع رکورد وجود دارد:

الف، رکوردهای پیش‌رونده مانند درازو نشست، انعطاف پذیری، پرش طول ،... هرچـه رکوردهـا بالاتر باشند بهتر هستند. برای محاسبه رکوردهای پیش رونده از فرمول زیر استفاده می‌کنیم.

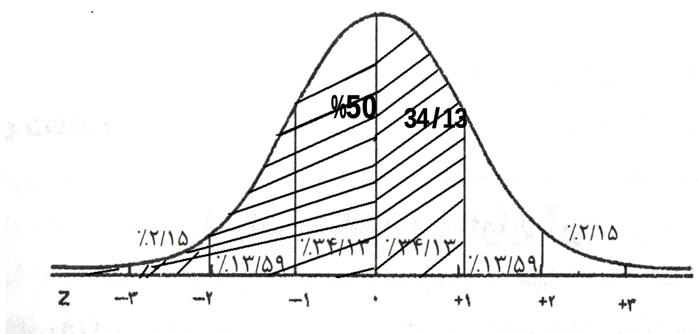
برای رکورد های پیش رونده $Z = \frac{X - \overline{X}}{S}$

ب، رکوردهای پس‌رونده مانند زمان دوی سرعت و دوهای مختلـف دیگـر، هرچـه زمـان رکـورد کمتر باشد بهتر است. درمحاسبه رکوردهای پس‌رونده علامت نمرات z عوض می‌شود ویا در فرمول z جای میانگین ونمره خام را عوض می‌کنیم.

برای رکوردهای پس رونده $z = \frac{\overline{x} - x}{s}$

مثال: اگر در دوی سرعت 45 متر رکورد فردی 6 ثانیه و میانگین و انحراف استاندارد به ترتیب 8 ثانیه و 2 ثانیه باشد آنگاه وی از چند درصد افراد بالاتر قرار دارد؟

برای رکوردهای پس رونده $z = \frac{\overline{x} - x}{s} \Rightarrow \frac{8-6}{2} = 1$



$$\%50 + \%34/13 = \%84/13$$

پس رکورد 6 ثانیه از 84/13 درصد افراد بالاتر قرار دارد.

نمره استاندارد T

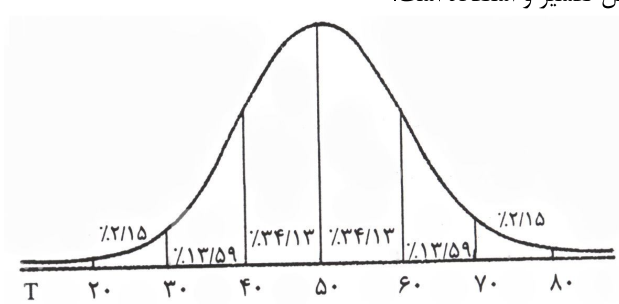
نمرات استاندارد دیگری که به نمرات T شهرت دارند، برای رفع مشکلات نمرات Z به کار می‌روند. مقیاس نمرات T، با ضرب Z در 10 (جهت حذف اعشار) و جمع‌کردن نمرات حاصل با 50 (جهت حذف نمرات منفی) به دست می‌آیید.

$T = 10z + 50$ یا $T = 10\left(\frac{x - \bar{x}}{s}\right) + 50$ برای رکوردهای پیش رونده

$T = 10\left(\frac{\bar{x} - x}{s}\right) + 50$ برای رکوردهای پس رونده

تذکر: در رکوردهای پس‌رونده برای به دست آوردن T جای نمره و میانگین عوض می‌شود.

بنابراین نمرات T دارای میانگین 50 و انحراف استاندارد 10 هستند. هرنمره T به تنهایی و به طور مستقیم قابل تفسیر و استفاده است.



شکل 2-5. نمودار توزیع طبیعی و نمره T

مثال: در صورتی که نمره هوش علی 71/5، نمره میانگین‌کلاس 67 و انحراف معیار (انحراف استاندارد) نمرات این کلاس 3 باشد، نمره T او را به‌دست آورید؟

حل :

$$T = 10\left(\frac{71/5 - 67}{3}\right) + 50 = 50 + 10(1/5) = 50 + 15 = 65$$

مثال: اگر در دوی سرعت 45 متر، رکورد فردی 8 ثانیه، میانگین آن 10 وانحراف آن 2 ثانیه باشد نمره T او را به‌دست آورید؟

برای رکوردهای پس رونده $T = 10\left(\frac{\bar{x} - x}{s}\right) + 50 \Rightarrow T = 10\left(\frac{10 - 8}{2}\right) + 50 \Rightarrow T = 60$

مثال: اگر دامنه نمرات z (± 3) در نظر گرفته شود نمرات T همین نمرات در چه دامنه‌ای قرار می‌گیرد؟

حل:

$T = 10z + 50$

$T = 10(+3) + 50 = 80$

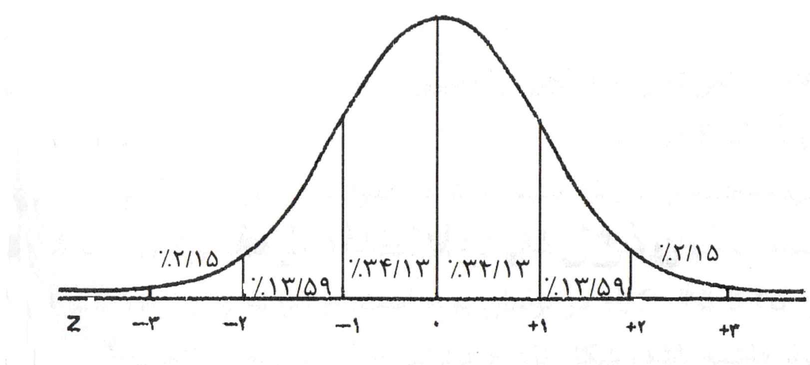
$T = 10(-3) + 50 = 20$

پس $20 \le T \le 80$ قرار می‌گیرد.

محاسبه مرتبه‌های درصدی

الف. محاسبه مرتبه درصدی در نمرات دسته‌بندی نشده:

مرتبه درصدی هر نمره خام عبارت است از: تعیین حد درصدی که درصد معینی از نمره زیر و بقیه بالاتر ازآن قرار می‌گیرند. برای محاسبه مرتبه درصدی، ابتدا با استفاده از فرمول z نمرات خام را به نمره استاندارد Z تبدیل کرده و سپس با استفاده از جدول سطوح زیر منحنی طبیعی، نمره‌ی z را به هم ارزش‌های زیر منحنی تبدیل می‌کنیم.

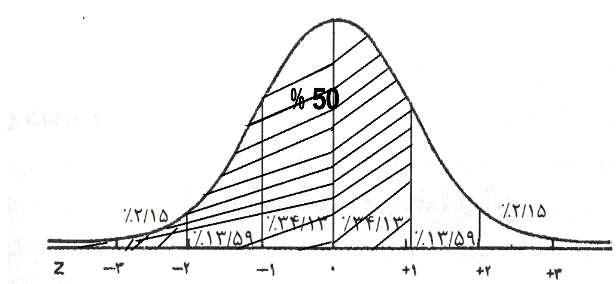


شکل 7-5: نمرات استادارد z

مثال: مرتبه درصدی فوتبالیستی که رکورد دراز و نشست وی 54 و میانگین رکورد آن‌ها 42 و انحراف استاندارد 12 باشد را به‌دست آورید؟

حل :

$$z=\frac{x-\bar{x}}{s}=\frac{54-42}{12}=\frac{12}{12}\Rightarrow z=1$$



$$\%50+\%34/13=\%84/3$$

مرتبه درصدی یک با توجه به منحنی طبیعی 13 /84 درصد می‌شود.

ب، محاسبه مرتبه درصدی در نمرات دسته‌بندی شده:

مرتبه‌های درصدی برای محاسبه‌ی درصد نمرات خام کاربرد دارد و جایگاه رکوردها و نمرات را در گروه به‌صورت درصد بیان می‌کند. در این روش، برای محاسبهٔ مرتبه‌های درصدی از فرمول زیر استفاده می‌شود[1]:

۱. PR: Percentile Ranking

$$P_R^1 = \frac{f(x-L)+cf(i)}{N(i)}$$

P = مرتبه‌ی درصدی   X = نمره‌ی خام

L = حد پایینی که نمره خام در آن طبقه قرار دارد.

Cf = فروانی تراکمی یک طبقه پایین‌تر از طبقه ای که نمره‌ی خام در آن واقع شده است.

N = تعداد اعداد

F = فراوانی طبقه ای که نمره خام در آ ن واقع شده است.

i = فاصله طبقات

مثال: درجدول توزیع فراوانی زیر، مرتبه درصدی نمره 32 را به‌دست آورید.

طبقات	F	Cf
35-39	1	20
30-34	3	19
25-29	7	16
20-24	5	9
15 – 19	3	4
10-14	1	1

L = ۲۹/۵

N =۲۰

i = ۵

x = ۳۲

cf =۱٦

توجه کنید که نمره 32 در طبقه 34-30 قرار دارد، لذا داریم:

$$P_R = \frac{f(x-L)+cf(i)}{N(i)} = \frac{3(32-29/5)+16(5)}{20(5)} = \frac{7/5+80}{100} = \frac{87/5}{100} \Rightarrow P_R = \%87/5$$

محاسبه نقاط درصدی

الف. محاسبه نقاط درصدی در نمرات دسته بندی نشده

در این مرحله نیز از فرمول نمرات استاندارد z استفاده مـی‌کنـیم، بـا ایـن تفـاوت کـه نقـاط درصدی یا مرتبه درصدی مشخص است و ملاک، نمره‌ها یا رکورد فرد می‌باشد. برای محاسبه رکورد یا نمره، ابتدا درصد مورد نظر را با استفاده از جدول سـطوح زیـر منحنـی طبیعـی بـه نمره‌ی استاندارد z تبدیل می‌کنیم (یعنی روش عکـس حالـت قبـل) و سـپس بـا اسـتفاده از فرمول نمرات استاندارد z، نمره خام را به‌دست می‌آوریم.

مثال: آزمونی با میانگین 60 و انحراف استاندارد 12 انجام شده است. مطلوب است نمره‌ی خام فردی که از 13/ 84 درصد از افراد گروه خود بالاتر قرار گرفته است؟

برای حل این مسئله، ابتدا با استفاده از جدول سطوح زیر منحنی طبیعی نمره z هم ارزش 84/13 درصد را مشخص می‌کنیم که برابر است با $z = +1$ سپس با استفاده از فرمول، نمره‌ی خام x را به‌دست می‌آوریم.

$$z = \frac{x - \bar{x}}{s} \Rightarrow 1 = \frac{x - 60}{12} \Rightarrow 12 = x - 60 \Rightarrow x = 72$$

ب. محاسبه نقاط درصدی در اعداد دسته بندی شده

نقاط درصدی زمانی مطرح می‌شود که درصد مشخص باشد و نمره خام، مورد محاسبه قرار گیرد. برای محاسبه‌ی نقاط درصدی از فرمول زیر استفاده می‌شود.

$$P_n^1 = L + \left(\frac{P_N - cf}{f}\right) \times i$$

P = نقطه درصدی

n = اندیس p و نمایش درصد افرادی است که در زیر صدک p_n ام قرار دارد.

L = حد پایین طبقه‌ای صدک مورد نظر در آن قرار دارد.

Cf = مجموع فراوانی‌های (فراوانی تراکمی) قبل از طبقه مورد نظر.

F = فراوانی ساده طبقه‌ای که، صدک مورد نظر در آن قرار دارد.

مثال: در درس حرکت شناسی نمره‌های زیر، توسط 20 دانشجو به دست آمده است. در صورتی که بخواهیم به 25 درصد دانشجویان جایزه‌بدهیم، برای اینکه فردی بتواند جایزه بگیرد باید چه نمره‌ای بگیرد؟

طبقات	F	Cf
19-20	2	20
17-18	3	18
15-16	5	15
13-14	6	10
11– 12	4	4
	N=20	

حل: برای انتخاب 25 درصد افراد جهت جایزه دادن باید 25 درصد بالاترین افراد را انتخاب کنیم.

پس 75 = 25 ــ 100، نقطه 75 درصدی را پیدا می‌کنیم.

$$P_{75}=\frac{75}{100}\times 20=15$$

75 درصدی در طبقه 16-15 قرار می‌گیرد.

$$P_{75}=14/5+(\frac{15-10}{5})\times 2\Rightarrow P_{75}=14/5+2\Rightarrow P_{75}=16/5$$

افرادی که حداقل، نمره 16/5 گرفته‌اند جایزه می‌گیرند.

مثال: در درس زبان تخصصی نتایج زیر به دست آمده، در صورتی که بخواهیم به15 درصد افراد، بورس تحصیلی اعطا کنیم حداقل نمره‌ای که فرد باید بگیردتا بـورس بـه او تعلـق بگیردچیست؟

طبقات نمرات	F	CF
18-20	3	20
15-17	3	17
12-14	2	14
9-11	4	12
6-8	5	8
3-5	3	3
0-2	0	0
	N=20	

حل: برای محاسبه نقاط درصدی در اعداد طبقه بندی شده بایـد، فراوانـی تجمعـی (تراکمـی cf) را بـه دست آوریم. برای این کار فراوانی مطلق هر طبقه را با فراوانی مطلق پایین جمع می‌کنیم و سـپس بـرای دست آوردن تعداد 15 درصد، باید نقطه 85 درصدی را بیابیم. چـون بهتـرین‌هـا در بـالای جـدول قـرار می‌گیرند.

درصدی $85=15-100$

$$P_n=L+(\frac{P_N-cf}{f})\times i$$

$$P_{85}=\frac{85}{100}\times 20=17$$ نفر

85 درصدی در طبقه 17- 15 قرار می‌گیرد.

$$P_{85}=14/5+(\frac{17-14}{3})\times 3=14/5+(\frac{3}{3})\times 3\Rightarrow p_{85}=17/5$$

باید حداقل نمره 17/5 بگیرند تا بورس تحصیلی به آن‌ها تعلق بگیرد.

مثال: رکورد 50 نفر در دوی 400 متر با مانع در جدول زیر آمده است، اگر بخواهیم 30 % افـراد را برای مسابقات آسیایی انتخاب کنیم حداقل رکورد انتخابی کدام است؟

زمان به ثانیه	فراوانی	CF
80-82	5	50
77-79	5	45
74-76	10	40
71-73	10	30
68-70	15	20
65-67	5	5
	N =50	

حل: قبل از حل مسئله، فراوانی تراکمی (cf) را محاسبه می‌کنیم (از قبل آن را حساب کردیم و در جدول قرار داده‌ایم). باید به این نکته مهم توجه کرد که، در این مسئله چون زمان مطرح شده است و ما می‌خواهیم کمترین زمان‌ها(رکوردها) را به مسابقات آسیایی اعزام کنیم پس ما فقط 30 درصد افراد پایین طبقه را انتخاب می‌کنیم که کمترین زمان نسبت به دیگران دارند.

$P_{30} = \frac{30}{100} \times 50 = 15$ نفر

پس 30 درصد افراد که تعداد آن 15 نفر می‌شود در طبقه 70 – 68 ثانیه قرار می‌گیرد.

$$P_{30} = L + (\frac{P_N - cf}{f}) \times i \Rightarrow P_{30} = 67/5 + (\frac{15-5}{15}) \times 3 \Rightarrow P_{30} = 67/5 + (\frac{30}{15})$$

$$\Rightarrow P_{30} = 69/5$$

حداقل رکوردی که برای اعزام مسابقات آسیایی لازم است 69/5 ثانیه می‌باشد.

فصل ۳

همبستگی

مقدمه

همبستگی عبارت است از، نوعی رابطه که ممکن است بین متغیرهای مختلف وجود داشته باشد. هدف از مطالعه همبستگی، شناخت و پیش داوری در مورد تأثیر احتمالی دو یا چند متغیر بر یکدیگر می‌باشد. در این فصل، هدف ما بررسی و مطالعه روابطی است که بین متغیرهای مختلف وجود دارد.

انواع همبستگی

اگر بین متغیرها، همبستگی وجود داشته باشد این همبستگی ممکن است به یکی از دو شکل زیر باشد:

1ـ همبستگی مستقیم (مثبت)

2ـ همبستگی معکوس (منفی)

اگر تغییرات اندازه‌های عددی دو متغیر، از نظر ریاضی در یک جهت باشدیعنی اگر یکی از متغیرها کم شود، متغیر دیگر نیز کم گردد و یا چنانچه یکی از آن‌ها افزایش یابد ، دیگری نیز افزایش یابد، (رابطه بین قد و وزن افراد در دوران رشد) می‌گوییم **همبستگی مستقیم** است و چنانچه افزایش یکی، متناظر با کاهش دیگر باشد می‌گوییم همبستگی معکوس است. رابطه بین ارتفاع کوه و فشار هوا و یا افزایش وزن و سرعت دویدن نمونه‌ای از **همبستگی معکوس** است.

نکته : همگنی واریانس بر ضریب همبستگی تأثیر زیادی دارد. همبستگی بین دو متغیر در جامعه ای که بر اساس متغیرهای مورد بررسی، ناهمگن است بیشتر از همبستگی همان متغیرها در جامعه‌ای است که بر حسب آن متغیرها همگن است.

شدت و ضعف همبستگی

اگر بین اندازه‌های دو متغیر مورد مطالعه همبستگی به‌گونه‌ای باشد که هر تغییری در انداره متغیر اول باعث تغییری متناسب در متغیر دوم شود (مستقیم یا معکوس) می‌گوییم همبستگی کامل

است، مانند رابطه بین قطر دایره و محـیط آن، امـا اگـر بـین دو متغیـر همبسـتگی کامـل نباشـد، همبستگی را ناقص می‌نامیم.

ابزارهای تشخیص همبستگی

برای تشخیص نوع همبستگی و مقدار شدت و ضعف آن می‌توان از ابزارهای مختلفی اسـتفاده کـرد که در این بخش به بررسی آنها خواهیم پرداخت.

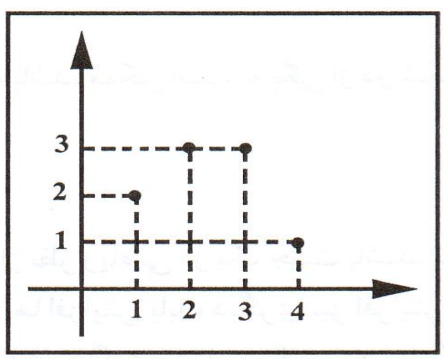
الف. نمودارهای پراکندگی

یکی از ساده‌ترین روش‌های بررسی وجود یا عدم وجـود همبسـتگی بـین دو متغیـر، رسـم نمـودار پراکندگی یا دیاگرام است. برای رسم چنین نموداری کافی است اندازه یکی از صفات را روی محـور طول‌ها (**ox**) و اندازه صفت دیگر را روی محور عرض‌ها (**oy**) قرار داده و برای هر فرد یک نقطه بـه دست آوریم که طول آن متناظر با اندازه صفت **x** و عرض آن با اندازه صفت **y** است.

مثال: اندازه‌های دو صفت x و y در یک جامعه آماری مطابق جدول زیر است :

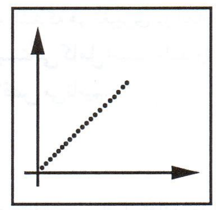
X	1	3	4	2
Y	2	3	1	3

نمودار پراکندگی متناظر با آن را رسم کنید.

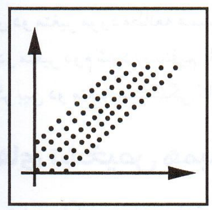


نکته‌ها:

1. اگر نقاط نمودار پراکندگی از پایین در سمت چپ و بالا در سمت راست باشد، مـی‌تـوان نتیجـه گرفت که همبستگی مستقیم است و چنانچه این نقاط در مسیر یک خط مستقیم باشد، هبسـتگی کامل است.

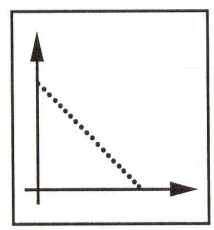


وجود همبستگی کامل و مستقیم

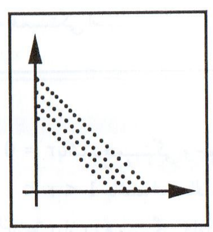


وجود همبستگی مستقیم

2. اگر حرکت نقاط دیاگرام از بالای سمت چپ به پایین سمت راست محورهای مختصات باشد همبستگی معکوس است و چنانچه نقاط دیاگرام ضمن داشتن این شرایط، در مسیر یک خط مستقیم نیز باشند، همبستگی از نوع کامل خواهد بود.

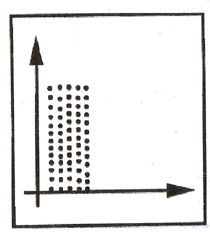


همبستگی معکوس و کامل

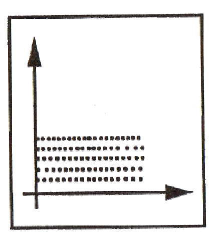


همبستگی معکوس

3. در مواردی که نقاط دیاگرام موازی محور طول‌ها یا عرض‌ها باشد و با هیچ نظم و قاعده‌ای روی محور مختصات قرار گرفته باشد ، می‌توان به نبودن همبستگی پی برد.



عدم همبستگی



عدم همبستگی

ب. ضریب همبستگی

رابطه همبستگی، شدت و جهت تغییرات دو متغیر را نسبت به یکدیگر نشان می‌دهد. منظور از شدت، مقدار کمی ضریب همبستگی[1] است که بین صفرتا ۱± در نوسان است. بدیهی است که هر

۱ . correlation coefficient

چه این مقدار به 1± نزدیکتر شود شدت همبستگی بین دو متغیر بیشتر خواهد بود. منظور از جهت، همسویی و یا عدم روند تغییرات دو متغیر با یکدیگر است. یک عدد مثبت معرف همسویی تغییرات بین دو متغیر، ویک عدد منفی، معرف تغییر دو متغیر در جهت مخالف یکدیگر است. به عنوان مثال افزایش میزان تغذیه با افزایش میزان وزن بدن، افزایش تمرین با مقدار توانایی بیشتر، افزایش سن تقویمی در دوران رشد با اضافه شدن وزن و ... همسو هستند و لذا ضریب همبستگی بین آنها مثبت خواهد بود. افزایش خستگی با میزان کارایی جسمانی، مقدار تمرین با رکورد سرعت، افزایش سن تقویمی در بزرگسالی با توانایی جسمانی، افزایش سختی کار با میزان انگیزش و ... در خلاف جهت یکدیگر حرکت می‌کنند و لذا همسبتگی، بین آنها منفی است. در مطالعات همبستگی رابطه بین دو صفت مورد بررسی قرار می‌گیرد.

محاسبهٔ ضریب همبستگی دراعداد خام

برای محاسبه‌ی همبستگی در اعداد خام چندین روش وجود دارد، که فقط به سه مورد از آن‌ها اشاره می‌شود.

1. استفاده از انحراف نمره ها از میانگین

$$r = \frac{\Sigma(x-\overline{x})(y-\overline{y})}{\sqrt{\Sigma(x-\overline{x})^2\Sigma(y-\overline{y})^2}}$$

مثال: در جدول زیر، نمرات بارفیکس و دراز و نشست 10 دانشجوی تربیت‌بدنی دانشکده فنـی و حرفه‌ای محمودآباد می‌باشد و نحوه محاسبه‌ی ضریب همبستگی دو دسته داده به شرح زیر است:

جدول 1-6. محاسبه ضریب همبستگی به روش انحرافات از میانگین را نشان می‌دهد.

X	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
20	12	7	2	49	4	14
18	16	5	6	25	36	30
16	10	3	0	9	0	0
15	14	2	4	4	16	8
14	12	1	2	1	4	2
12	10	-1	0	1	0	0
12	9	-1	-1	1	1	1
10	8	-3	-2	9	4	6
8	7	-5	-3	25	9	15
5	2	-8	-8	64	64	64
$\Sigma x = 130$ $\bar{x} = 13$	$\Sigma y = 100$ $\bar{y} = 10$	0	0	$\Sigma(x-\bar{x})^2$ $=188$	$\Sigma(y-\bar{y})^2$ $=138$	$\Sigma(x-\bar{x})(y-\bar{y})$ $=140$

$$r = \frac{140}{\sqrt{18\,8(138)}} = 0/87$$

2. محاسبه‌ی ضریب همسبتگی گشتاوری پیرسون

روش پیرسون[1] زمانی بکار برده می‌شود که داده‌های جمـع‌آوری شـده بـا اسـتفاده از مقیـاس فاصله‌ای یا نسبی اندازه‌گیری شده باشد.

برای محاسبه‌ی ضریب همبستگی از فرمول زیر استفاده می‌کنیم. در تربیت‌بدنی بیشـتر بـرای تست رشته بسکتبال به‌کار می‌رود.

$$r = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2]\ [N\Sigma y^2 - (\Sigma y)^2]}}$$

۱. Pearson

مثال: اعداد متغیرهـای x و y در دسـت اسـت. ضـریب همبسـتگی آن را از طریـق گشـتاوری پیرسیون محاسبه نمائید.

متغیر X	متغیر y	X^2	y^2	xy
24	13	567	169	312
20	9	400	81	180
18	12	324	144	216
17	20	289	400	340
15	11	225	121	165
12	16	144	256	192
10	5	100	25	50
8	2	6	4	16
6	7	36	49	42
4	1	16	1	4
$\sum x = 134$	$\sum y = 96$	$\sum x^2 = 2174$	$\sum y^2 = 125$	$\sum xy = 1517$

$$r = \frac{10(1517) - (134)(96)}{[10(2174) - (134)^2]\,[10(1250) - (96)^2]} \Rightarrow r = \frac{2306}{\sqrt{12426656}}$$

$$\Rightarrow r = \frac{2306}{3525/6} \Rightarrow r = 0/65$$