آمار $\mathcal{DLM}$ شامل ۱۸ سرفصل درقالب ۲۲۱۷ فلش کارت + کتابچه CTS می‌باشد. (زمان بندی مطالعه ۸۴ روز)

اکیدا به همه داوطلبان کارشناسی ارشد و دکتری مدیریت، حسابداری و اقتصاد توصیه می‌شود.

روش مطالعه آمار CTS می باشد که یک روش منحصربفرد و بسیار مدرن برای مطالعه دروس ریاضی-محور است.(توضیحات در ادامه آمده)

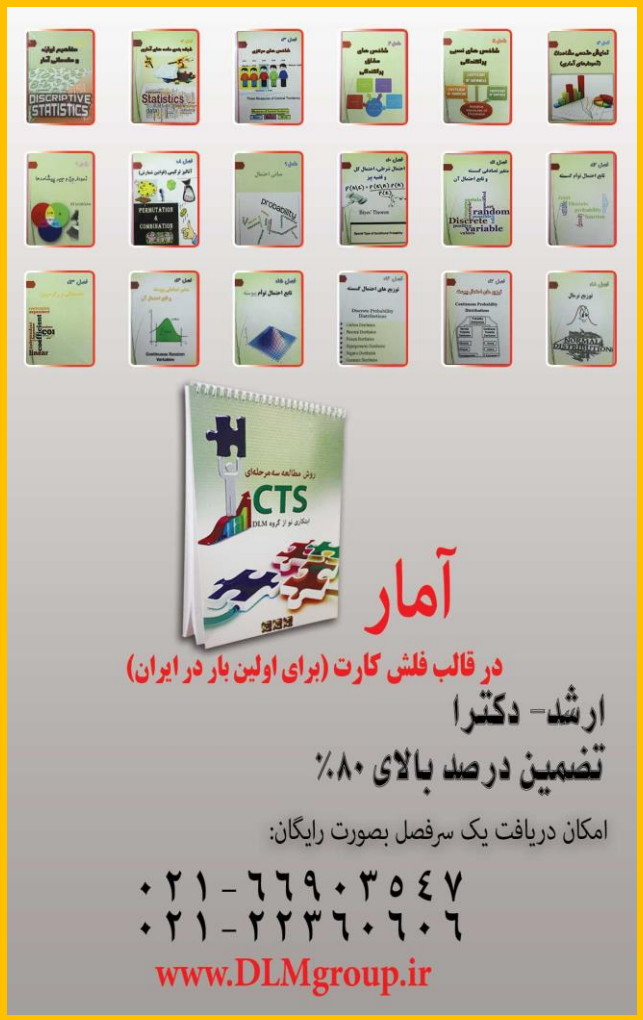
CTS چیست؟

یک کتابچه ۶۹ صفحه‌ای است که ما در آن به دقت برای شما استخراج کرده‌ایم که هر روز از مطالعه فلش کارتها چه مطالبی را می‌بایست آموخته باشید و در قالب جای خالی در اختیار شما قرار داده‌ایم. تعاریف، مفاهیم‌ها، فرمول‌ها و ... را از فلش‌کارت‌ها خوانده و در CTS خلاصه برداری می‌کنید (یعنی جاهای خالی را که ما در اختیار شما قرار داده‌ایم پُر می‌کنید) و شماره مسئله‌ها و تست‌هایی هم که باید حل کنید به دقت تعیین شده؛ در نهایت کل ۲۲۱۷ فلش‌کارت تبدیل می‌شوند به یک کتابچه ۶۹ صفحه‌ای که پایه آمار شما را برای همیشه خواهد ساخت و در تمام زندگی علمی خود می‌توانید مرتب بدان مراجعه کنید.

[مرور مطالب (براساس برنامه زمان‌بندی مطالعه که در انتهای کتابچه CTS در اختیار شما قرار داده شده از کتابچه CTS انجام می‌شود.]

در این فایل نیمی از فلش‌کارت‌های فصل ۳ + آن قسمت از کتابچه CTS که مربوط به این فلش‌کارت‌هاست تقدیم شده است. ضمناً کل کتابچه CTS هم در این فولدر قرار داده شده تا بتوانید فهرست تمام مطالب آموزش داده شده در کل پک را در اختیار داشته باشید. اگر هم حوصله پرینت و برش فلش‌کارت‌ها یا مطالعه از روی تبلت و گوشی را ندارید با ۰۲۱۶۶۹۰۳۵۴۷ تماس حاصل فرمایید تا سمپل رایگان، چاپ‌شده و حاضر و آماده در اختیار شما قرار گیرد.

DISCRIPTIVE STATISTICS
Statistics
Three Measures of Central Tendency
Relative Measures of Dispersion
Bayes' Theorem
PERMUTATION & COMBINATION
probability
random
Discrete variable
Continuous Random Variables
Discrete Probability Distributions
Continuous Probability Distributions
NORMAL DISTRIBUTION


CTS
آمار
در قالب فلش کارت (برای اولین بار در ایران)
ارشد- دکترا
تضمین درصد بالای ۸۰٪
امکان دریافت یک سرفصل بصورت رایگان:
۰۲۱-۶۶۹۰۳۵۴۷
۰۲۱-۲۲۳۶۰۶۰۶
www.DLMgroup.ir

منابعی که پک آمار DLM پوشش می‌دهد:

۱. **آمار و احتمال، آموزشگاه پارسه، محسن طوراني**
۲. **آمار و احتمالات، آموزشگاه مدرسان شریف**
۳. **آمار و احتمال، هادي رنجبران**
٤. **آمار کاربردي ۱، عالِم تبریز، انتشارات پوران پژوهش**
٥. **آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار**
٦. **آمار و کاربرد آن در مدیریت، عادل آذر و منصور مؤمني**
۷. **آزمون‌های تألیفي آمار کاربردي، وحید انصاري**
۸. **آمار و کاربرد آن در مدیریت، انتشارات قلم چي**
۹. **آمار، آموزشگاه ماهان**
۱۰. **آمار، آموزشگاه علوي**

سایر منابع مورداستفاده:

۱. آمار: روش‌ها و کاربردها، حمیدی زاده
۲. آمار و کاربرد آن در مدیریت، صفاری
۳. جزوه آمار، مهرداد پرچ، جهاد دانشگاهی دانشگاه تهران و مطالبی از وبسایت ایشان
٤. اصول، مبانی و مفاهیم علم آمار، علی بشارت
٥. آمار کاربردی، علی عمیدی
٦. مفاهیم و روش‌های آماری، شهرآشوب
۷. اصول و روش‌های آماری، فرشاد فر
۸. روش‌های آماري، پرویز تاجداري
۹. کاربرد آمار در مدیریت ترافیک
۱۰. (مجموعه سؤالات آمار و روش تحقیق، صمدی
۱۱. ۱۰۰۰ مسئله حل‌شده در آمار و احتمالات مهندسی، اکبری و تراب زاده
۱۲. جبر و احتمال، پرسمان گاج
۱۳. ریاضیات گسسته، ۱۰ استاد
۱٤. ریاضیات گسسته، ۸ کتاب ریاضی گاج
۱٥. ریاضیات گسسته، گاچ

کارشناسی ارشـــــــد و دکتری فقط با DLM میچسبه! و قبولی هم که قطعیه! پس دیگه چی میخواین؟!!

روش مطالعه ۳ مرحله ای (CTS)

Dynamic Learning Method

توضیح: فلش کارت های آمار شامل دو دسته «**حفظ کردنی**» و «**حل کردنی**» می‌باشند. «حفظ کردنی» مانند تعاریف و مفاهیم و «حل کردنی» مانند مثال ها و تست ها.

(تعریف مطالعه مقدماتی، مطالعه نیم چاشت و مطالعه نهایی)

مطالعه اول (مقدماتی)	۱- کل فلش‌کارت ها را یک دور مطالعه می‌کنیم؛ شامل فلش‌کارت‌های حفظی(مانند تعاریف) و فلش کارت های حل کردنی (مانند مثال ها و تست ها) ۲- جواب مسئله ها و تست ها را نگاه می کنیم (در این مرحله هنوز لازم نیست خودتان آنها را بر روی کاغذ سفید حل کنید؛ کافی است روش حل را کاملا درک کنید.) **نکته:** در این مرحله مهم این است که یاد بگیرید داده های مساله چگونه در فرمول جایگذاری شده‌اند (یا چگونه از آنها استفاده شده است.)
مطالعه دوم (نیم چاشت)	۳- حالا مجددا از نو، فلش کارت ها را با دقت بیشتری مطالعه می کنیم. *حفظی* ها را کافیست یکبار دیگر از رو بخوانید و مفهوم آنها را با زبان خودتان برای خود بازگو کنید. در مورد فلش کارت های *حل کردنی* باید سعی کنید راه حل مسئله‌ها، مثال‌ها و تست‌ها را کاملا یاد بگیرید؛ یعنی صورت مساله را بخوانید. قلم دست بگیرید و بر روی کاغذ سفید حل شان کنید. در این مرحله، هر وقت لازم شد می‌توانید به راه حل نگاه کنید. (مثلا حل مثال را آغاز می کنید و تا اواسط فرایند حل پیش می‌روید. ناگهان در این مرحله یک چیزی را فراموش می کنید. به فلش کارت مراجعه می کنید، آن قسمتش را نگاه می کنید، برایتان یادآوری می شود و دوباره حلتان را ادامه می دهید و اگر دوباره به جایی رسیدید که لازم شد به حل فلش کارت مراجعه کنید، این کار را انجام می دهید.)
مطالعه سوم (نهایی)	۴- به **باکس CTS** (در صفحه بعدی این راهنما) مراجعه کنید. مفاهیم مهم، فرمول‌ها و مسئله‌هایی که باید یاد بگیرید، در باکس CTS استخراج شده‌اند. * **مسئله‌ها**یی را که شماره فلش‌کارت آنها را مشخص کرده‌ایم بر روی کاغذ سفید حل می‌کنید. (اینبار بدون نگاه کردن به فلش‌کارت) * **فرمول**ها را در جاهایی که بدین منظور پیش‌بینی شده وارد می‌کنید. (در این مرحله بهتر است فرمول‌ها را ابتدا به خاطر بسپارید و از حفظ وارد باکس CTS کنید و یکبار هم با فلش‌کارتی که آن فرمول در آن فلش‌کارت قرار دارد مطابقت دهید که مطمئن شوید فرمول کاملاً درست و دقیق وارد باکس CTS شده است.) * **مفاهیم** را به زبان خودتان در محل‌های تعیین شده وارد نمایید. * **جاهای خالی(نقطه‌چین)**ها را پُر کنید. (اگر مرحله قبلی (مطالعه نیم‌چاشت) را به درستی انجام داده باشید برای پُر کردن *جای خالی‌ها*، *مفاهیم*، *فرمول‌ها* و نیز حل *مسئله‌ها و مثال‌ها* نیازی به مراجعه به فلش‌کارت‌ها نخواهید داشت. اگر هم احیاناً نیاز شد مراجعه کنید، اشکالی ندارد. در ادامه مطالب آنقدر مرور و تکرار خواهند شد که خود به خود در ذهن شما جای بگیرند. اما اصل روش این است که باکس CTS را *از حفظ* پُر کنید.)

- **باکس CTS چیست؟**

همانطوریکه احتمالاً تاکنون متوجه شده‌اید «باکس CTS» در واقع مفاهیم، نکات، فرمول‌ها، مثال‌ها و تست‌هایی است که پس از هر روز مطالعه (مطابق با جدول زمان‌بندی) باید یاد گرفته باشید. در پایان مطالعه پک، یک کتابچه منحصربفرد در اختیار خواهید داشت (جدول CTS) که تمامی مطالب آمار۱ را در بر می‌گیرد و هر وقت لازم بود می‌توانید بدان مراجعه نمایید.

به این ترتیب فرمول‌ها و مفاهیم **همه یکجا جمع‌بندی** شده‌اند و هر وقت نیاز داشته باشید آنها را مرور کنید لازم نیست بروید سراغ فلش‌کارت‌ها و کلی وقت صرف پیدا کردن فرمول، تعریف یا مثال مورد نظرتان کنید.

ضمن اینکه چون مفاهیم را از حفظ و به زبان خودتان وارد باکس CTS می‌کنید، گویی دارید آنها را به شخص دیگری آموزش می‌دهید که این امر به معنای این است که مطلب را درک کرده‌اید. تست‌ها و مسئله‌ها را هم که باید بدون نگاه به فلش‌کارت‌ها بر روی کاغذ سفید حل کنید. خوب، چاره‌ای ندارید جز اینکه آنها را یاد گرفته باشید.

همچنین در مرورها نیازی به رجوع به فلش‌کارت‌ها (بجز در مواردی که مطالب را احیاناً فراموش کرده باشید و نیز در مورد مسائل و تست‌ها به منظور خواندن صورت مساله از روی فلش‌کارت) نخواهید داشت و **مستقیماً به باکس CTS رجوع خواهید کرد**. این کار صرفه‌جویی زمانی چشمگیری در پی خواهد داشت بدون آنکه حتی یک نکته را از دست بدهید.

- روش CTS با در نظر گرفتن مهم‌ترین مولفه‌های یادگیری منجمله «یادگیری فعال (Active Learning)» و در ادامه روش۵Thicks براساس مطالعه نیازها و نظرات داوطلبان و جهت‌گیری TQM توسط گروه DLM ابداع شده است.

این یک مخصوص رشته‌های مدیریت، حسابداری و اقتصاد (ارشد و دکترا) با لحاظ کردن تفاوتهای اندکی که در سوالات آمار این سه رشته وجود دارد طراحی گردیده اما سایر رشته‌ها نیز می‌توانند آنرا مطالعه کنند *(منابع پوشش داده شده را ببینید)*

به ضرس قاطع یک آمار DLM با فاصله زیادی نسبت به سایر منابع جامع‌ترین منبع فارسی (و حتی لاتین) آموزش « آمار ۱ » در بازار نشر ایران است. این را پس از مطالعه یک، تصدیق خواهید فرمود.

در کنکور ارشد و دکترا سوالی از آمار ۱ مطرح نخواهد شد که با مطالعه DLM نتوانید به آن پاسخ دهید. رشته‌ها و مقاطعی که آمار۲ را هم باید مطالعه کنند، ابتدا باید آمار ۱ را کاملاً مسلط باشند تا بتوانند آمار۲ را به خوبی درک کنند. پس مطالعه «آمار DLM» برای همه رشته‌ها و همه مقاطع توصیه می‌شود.

برای ارشد مدیریت، حسابداری و اقتصاد از مباحث آمار ۲ فقط یک یا دو سوال مطرح می‌شود. پس مطالعه آمار ۲ که حجمی تقریبا معادل آمار ۱ دارد اصلا منطقی نیست.

(کسانی که مایلند پس از مطالعه یک DLM آمار ۲ را هم بخوانند می‌توانند کتاب آمار ۲ دکتر عادل آذر(*زبان ساده‌تر*) و آمار۲ محسن طورانی[پارسه](*مطالب جامع تر*) را مطالعه کنند.)

*ما زمان لازم برای مطالعه یک «آمار» را ۸۳ روز اعلام می‌کنیم. جدول CTS ۱۶۵ را نشان می‌دهد. این، بدان دلیل است که مطالعه هر دو روز از جدول CTS در یک روز کاملاً روتین و منطقی است. اما ما جدول CTS را بر مبنای یک مطالعه آرام، ملایم و عمیق طراحی کرده‌ایم. آمار و ریاضی را می‌شود **با لذت** خواند. "چه کاریه اگر وقت دارید عجله کنید! بذارید همون ۱۶۵ روز طول بکشه!"*
اما اگر یک را دیر تهیه کرده‌اید و زمان کمتری تا کنکور باقیست، دو روز (یا ۳ روز) را هم به راحتی می‌توانید در یک روز بخوانید.

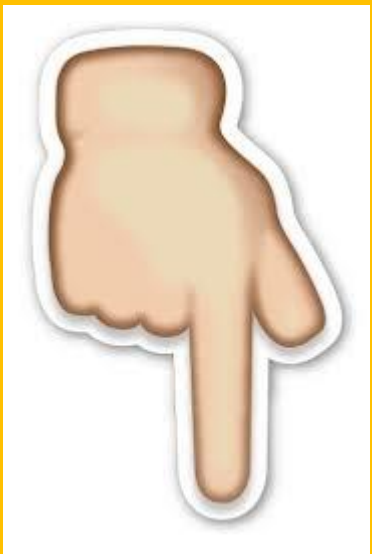
روز	**باکس CTS (فصل ۳):** « **جمع بندی خلاصه نکات ، فرمول‌ها و روابط مهم هر سرفصل**»	فلش کارت
روز ۱۲	**نکته ۱:** شاخص مرکزی مُد (نما) **منحصر بفرد** .....،یعنی ..... اگر دو یا چند داده (**نه همه داده‌ها**) بیشترین تکرار را داشته باشند، اون‌وقت..... را به عنوان مد مشاهدات در نظر می‌گیریم. اگر **تمامی** داده‌ها (x_i ها) به یک اندازه تکرار شده باشند (یعنی **فراوانی مطلق یا نسبی یکسانی** داشته باشند)، آنگاه مد جامعه ..... است.	۱۷۲ ۱۷۲
روز ۱۲	**خاصیت مهم مد (با مثال):** اگر از **مد** به عنوان نماینده(برای تخمین مقدار عددی تک تک مشاهدات) استفاده کنیم، آنگاه ..... **حداقل خواهد بود.**	۱۷۶
روز۱۲	**حل فیشهای : ۱۷۱-۱۷۳-۱۷۴-۱۷۶**	
روز ۱۳	مراحل محاسبه مد در داده های نوع سوم (۳ مرحله): الف) ب) ج) d_1 : ..... : $\boxed{d_1 = ...}$ یا $\boxed{d_1 = ...}$ d_2 : ..... : $\boxed{d_2 = ...}$ یا $\boxed{d_2 = ...}$ مقدار مدی که از فرمول فوق بدست می‌آوریم، **باید حتماً ..... قرار داشته باشد.** **نکته ۳)** اگر **طبقه اول** جدول، طبقه مد دار باشد: $d_1 = ...$ و اگر **طبقه آخر**، طبقه مد دار باشد: $d_2 = ...$	۱۷۷ ۱۷۸ ۱۷۸ ۱۸۱ ۱۸۲
روز ۱۳	**روش تستی برای محاسبه مد:** *روش معمولی و کم دقت برای محاسبه مد در جداول طبقه‌بندی شده (که به صورت (L-U هستن) اینه که .....* *رو به عنوان مد مشاهدات در نظر بگیریم.اما مقدار واقعی و دقیق مد، بلکه* **به سمت طبقهٔ مجاوری که .....متمایل میشه.**	۱۸۴
روز۱۳	**حل فیشهای : ۱۷۸-۱۸۲-۱۸۶**	
روز ۱۴	*در رسم نمودار بافت نگار:* *پر روی محور افقی، .... قرار می گیره.* *و پر روی محور عمودی هم، ..... قرار می گیره.*	۱۹۶
	پراکندگی و تغییرپذیری مد ..... است و در نتیجه **پایداری و ثبات آن .....**است، به همین علت **مد را** ..... **شاخص مرکزی**قلمداد می‌کنند.	۲۰۱
روز ۱۴	**کاربرد مد (نما)**:هر گاه در یک جامعه، ..... معیار سنجش و انتخاب ما با شه (مثلاً هنگام .....) اون وقت باید از شاخص مرکزی **مد (نما)** استفاده کنیم.	۲۰۳
	نکته بسیار مهم: *در بیان میانه، فقط می‌توان از عبارت زیر استفاده کرد:* **۱**- ۵۰ درصد مشاهدات**..... یا مساوی**میانه هستند (یعنی ۵۰ درصد مشاهدات، ..... برابر میانه هستند). **۲**- ۵۰ درصد مشاهدات **..... از** میانه هستند.	۲۰۵
	نکته بسیار مهم: شاخص میانه (**Md**) جزء ۵۰ درصد **مشاهدات ..... از خودش** است، بنابراین موقع تفسیر میانه، بجای علامت ..... باید حتماً از علامت..... استفاده کنیم.	۲۰۵
روز ۱۴	نحوه محاسبه میانه در داده های نوع اول: محل میانه = ... ⟵ اگه N**فرد**باشه محل میانه = ... ⟵ اگه N**زوج**باشه	۲۰۸
روز۱۴	**حل فیشهای : ۱۹۹-۲۰۳-۲۰۴**	
روز ۱۵	**نحوه محاسبه میانه در داده های نوع دوم با داشتن فراوانی مطلق (در ۴ گام):**	۲۱۱
	الف)مرتب کردن x_i ها (طبقات جدول) به صورت	
	ب)یافتن **محل** میانه: $C_{Md} = ...$:**محل میانه**	
	ج)محاسبه فراوانی **تجمعی** طبقات F_{C_i}: $F_{C_i} = ...$	

	یادآوری: برای محاسبه فراوانی **تجمعی** هر طبقه (F_{C_i})، فراوانی ..... آن طبقه را با فراوانی .....جمع می‌کنیم. **د)یافتن طبقه میانه دار:** *اولین طبقه جدول از چپ به راست ، که در آن.....باشد.*	
روز ۱۵	**نحوه محاسبه میانه در داده های نوع دوم با داشتن فراوانی نسبی (در ۳ گام):**	۲۲۱
	گام ۱) مرتب کردن طبقات (x_iها) به صورت	
	گام ۲) محاسبه فراوانی **تجمعی نسبی** طبقات (f_{C_i}): f_{C_i} = ...	
	گام ۳) یافتن طبقه میانه دار:اولین طبقه‌ای که فراوانی تجمعی نسبی آن ..... باشد.	
روز ۱۵	**نحوه محاسبه میانه در داده های نوع سوم با داشتن فراوانی مطلق (در ۳ گام):**	۲۲۳
	گام ۱)محاسبه فراوانی **تجمعی** طبقات: ..... = F_{C_i}	
	گام ۲)یافتن **محل میانه (طبقه میانه دار)**:اولین طبقه‌ای از چپ به راست که فراوانی **تجمعی**اش ..... باشد.	
	گام ۳)یافتن **مقدار** میانه از رابطه : = **میانه**	
	تذکر: اگر طبقات **گسسته** باشند، بعد از یافتن طبقه میانه دار، باید آن را ..... کنیم. یعنی باید حدود ..... طبقه میانه دار را بدست بیاریم و سپس گام (۳) را انجام بدیم.	۲۲۴
روز۱۵	**حل فیشهای : ۲۱۱-۲۱۳-۲۱۵-۲۱۷-۲۲۳**	
روز ۱۶	**نحوه محاسبه میانه در داده های نوع سوم با داشتن فراوانی نسبی (در ۳ گام):**	۲۳۱
	گام۱) محاسبه **فراوانی تجمعی نسبی** طبقات: f_{c_i} = ...	
	گام۲) یافتن محل میانه (طبقه میانه‌دار): اولین طبقه‌ای که در آن ..... است.	
	گام ۳)یافتن **مقدار** میانه از رابطه : = **میانه**	۲۳۱ ۲۳۲
	تذکر مهم:*در صورت گسسته بودن طبقات، قبل از انجام گام۳، اول باید طبقه میانه‌دار را ..... کنیم، یعنی باید حدود..... طبقه میانه‌دار را بدست آوریم.*	
	نکته: قبلاً در مورد مد گفتیم که مقدار مد باید حتماً بین ..... قرار داشته باشد. در مورد میانه هم می‌تونیم بگیم که مقدار میانه هم باید حتماً بین ..... باشد.	
	نکته تستی:*اگر در سطر فراوانی نسبی تجمعی، مستقیماً ..... رو ببینیم، اون وقت تنها کافیه که ..... رو به عنوان* **میانه** *در نظر بگیریم.*	۲۳۳
	خواص میانه: $\xrightarrow{y_i = ax_i} Md_y$ = ...	۲۳۶
	$\xrightarrow{y_i = x_i \pm b}$ Md_y = ... $\xrightarrow{y_i = ax_i \pm b}$ Md_y = ...	۲۳۷
روز۱۶	**حل فیشهای : ۲۲۵-۲۲۷-۲۳۱**	
روز ۱۷	**خاصیت مهم میانه:** مجموع ..... *انحرافات (تفاضلات) داده‌ها از میانه، ..... است:* $\sum_{i=1}^{k}$ .....=	۲۴۰
	نکته بسیار مهم در تست‌ها: باید توجه داشت که این خاصیت مهم میانه، همواره باید به صورت ..... بیان شود.	۲۴۲
روز ۱۷	**فاصله**گاراژ iأم تا پمپ بنزین: .....=	۲۴۵
روز ۱۷	کاربردهای میانه (۳مورد):	۲۵۵
روز۱۷	**حل فیشهای : ۲۴۴-۲۴۵-۲۴۸-۲۵۳-۲۵۵**	

روز ۱۸	**انواع چندکها (۳ مورد):** مفهوم چارک اول: مقداری که $\frac{۱}{۴}$ = ۲۵٪ مشاهدات ..... آن هستند و $\frac{۳}{۴}$ = ۷۵٪ مشاهدات .....از آن هستند. مفهوم چارک سوم: مقداری که $\frac{۳}{۴}$ = ۷۵٪ مشاهدات ..... آن هستند و $\frac{۱}{۴}$ = ۲۵٪ مشاهدات .....از آن هستند.	۲۵۷ ۲۵۸
روز ۱۸	**دهک**: اگر دامنهٔ داده‌های جامعه آماری را به ..... تقسیم کنیم، به طوری که هر یک از این قسمت‌ها ..... **یا** .....از کل فراوانی‌ها را دربرداشته باشند، آنگاه دهک‌های ..... **تا** ..... بوجود می‌آیند. دهک اول: مقداری که $\frac{۱}{۱۰}$ = ۱۰٪ مشاهدات ..... آن هستند و $\frac{۹}{۱۰}$ = ۹۰٪ مشاهدات .....از آن هستند.	۲۶۱ ۲۶۲
روز ۱۸	**صدک**:اگر دامنه داده‌های جامعه آماری را به ..... تقسیم کنیم، به طوری که هر یک از این قسمت‌ها، ..... از کل فراوانی‌ها را در برداشته باشد، آنگاه صدک‌های ..... **تا** .....بوجود میان.	۲۶۴
روز ۱۸	تطبیق چندکها با میانه: صدک .....= دهک .....= چارک .....= میانه	۲۶۶
روز ۱۸	**نحوه محاسبه چندکها در داده های خام (۳ گام):** **گام ۱)**مرتب کردن داده‌ها به صورت .....و کدگذاری اونا از ۱ تا N **گام۲)**یافتن **محل چندک** (چارک، دهک و صدک) مورد نظر از رابطه: **: محل چارک a ام** **: محل دهک a ام** **: محل صدک a ام** **گام۳)**یافتن **مقدار چندک:**	۲۶۸
روز ۱۸	**نکته**: برای یافتن مقدار ۸/۳اُمین داده باید اول دادهٔ .....رو در نظر بگیریم و بعد ۰/۳ فاصله بین ..... رو به دادهٔ ..... اضافه کنیم: $x_{۸/۳}$ = ...+.....	۲۶۹
روز ۱۸	**نحوه محاسبه چندکها در داده های نوع دوم با داشتن فراوانی مطلق (۳ گام):** **گام ۱)** .....کردن طبقات و محاسبه ..... طبقات. **گام۲) یافتن محل چندک**از رابطه: **محل چارک a ام** $= C_{Q_a}$ = **محل دهکa ام** $= C_{D_a}$ = **محل صدک a ام** $= C_{P_a}$ = **گام۳) یافتن مقدار چندک:**از چپ به راست اولین طبقه (x_i)ای را انتخاب می‌کنیم که: ۱) برای چارکها:..... $\geq F_{C_i}$ ۲) برای دهکها :..... $\geq F_{C_i}$ ۳) برای صدکها :..... $\geq F_{C_i}$	۲۷۱
روز۱۸	**حل فیشهای : ۲۵۹-۲۶۰-۲۶۳-۲۶۸-۲۷۰**	
روز ۱۹	$x_{۱۴/۹}$ = ...+.....	۲۷۵
	نحوه محاسبه چندکها در داده های نوع دوم با داشتن فراوانی نسبی (۲ گام): **گام ۱)** ..... کردن طبقات و محاسبهٔ .....طبقات **گام ۲)** یافتن محل چندک:اولین طبقه(x_i)ای که: ۱) برای **چارک a**اُم ..... $\geq f_{C_i}$ باشه. **مثال**: برای چارک سوم باید ..... $\geq f_{C_i}$ باشد. ۲) برای **دهک a**اُم ..... $\geq f_{C_i}$ باشه.**مثال**: برای دهک ششم ($a = ۶$)، باید: ..... $\geq f_{C_i}$ ۳) برای **صدک a**اُم ..... $\geq f_{C_i}$ باشه.**مثال**: برای صدک بیستم ($a = ۲۰$)، باید: ..... $\geq f_{C_i}$	۲۷۶

روز ۱۹	**نحوه محاسبه چندکها در داده های نوع سوم با داشتن فراوانی مطلق (۴ گام):** **گام ۱) محاسبه فراوانی .....** طبقات: **گام ۲) یافتن محل چندک (طبقه چندک دار):اولین** طبقه ای که فراوانی ..... بیشتر یا مساوی..... یا ..... یا ..... باشد،یعنی: (..... و ..... و ..... $\geq F_{C_i}$) **گام۳)..... کردن** طبقه **چندک دار** در صورت ..... بودن. **گام۴) یافتن مقدار تقریبی چندک** از روابط زیر: **چارک aام** $Q_a \approx$ ..... ؛ $a = ۱,۲,۳$ **دهک a ام** $D_a \approx$ ..... ؛ $a = ۱,۲,...,۹$ **صدک aام** $P_a \approx$ ..... ؛ $a = ۱,۲,...,۹۹$ **نکته:** *تمام محاسبات مربوط به پارامترهای مرکزی (مد، میانه، میانگین و چندک‌ها) با فرض ..... بودن حدود طبقات انجام می‌شوند، یعنی قبل از محاسبه این شاخصها، اول باید طبقات مون رو ..... کنیم.*	۲۷۸ ۲۷۹ ۲۸۲
روز۱۹	**حل فیشهای : ۲۷۳-۲۷۶-۲۸۱-۲۸۵**	
روز ۲۰	**نکته** *برای رسیدن به حد پائین طبقه چهارم، فقط کافیه به اندازهٔ .....به ..... اضافه کنیم:* $L_۴ = ... + ...$	۲۹۰
روز ۲۰	**نکته مهم:**اگر محل چندک **دقیقاً در ستون فراوانی تجمعی دیده شود**، در این شرایط می‌توان طبقهٔ .....یا طبقه ..... را به عنوان طبقه چندک‌دار در نظر گرفت و نتیجه هر دو حالت، یکسان خواهد بود. **نکته تستی**: اگه محل چندک (یعنی $\frac{aN}{۴}$، $\frac{aN}{۱۰}$ یا $\frac{aN}{۱۰۰}$) **رو دقیقاً تو بین مقادیر فراوانی تجمعی ببینیم، آنگاه بدون نیاز به هیچ محاسبه‌ای** می‌تونیم بگیم: **چندک مورد نظر ما برابر با ..... است.** اما یادمون باشه که: **فقط موقعی می‌تونیم از این راه حل تستی استفاده کنیم که**	۲۹۵ ۲۹۵ ۲۹۶
روز ۲۰	**نحوه محاسبه چندکها در داده های نوع سوم با داشتن فراوانی نسبی (۴ گام):** **گام ۱)** محاسبهٔ فراوانی .....طبقات **گام ۲)** یافتن **محل** چندک:اولین طبقه(x_i)ای که: ۱) برای **چارک aام** ..... $\geq f_{C_i}$ باشه. **مثال**: برای چارک سوم باید ..... $\geq f_{C_i}$ باشد. ۲) برای **دهک aام** ..... $\geq f_{C_i}$ باشه.**مثال**: برای دهک ششم ($a = ۶$)، باید: ..... $\geq f_{C_i}$ ۳) برای **صدک aام** ..... $\geq f_{C_i}$ باشه.**مثال**: برای صدک بیستم ($a = ۲۰$)، باید: ..... $\geq f_{C_i}$ **گام ۳)..... کردن**طبقات در صورت ..... بودن. **گام ۴) یافتن مقدار تقریبی چندک aام** (a: شماره چندک) از روابط زیر: **چارک aام** $Q_a \approx$ ..... ؛ $a = ۱,۲,۳$ **دهک a ام** $D_a \approx$ ..... ؛ $a = ۱,۲,...,۹$ **صدک aام** $P_a \approx$ ..... ؛ $a = ۱,۲,...,۹۹$ **نکته تستی**:*اگه طبقات ما **از همون اول، ..... باشن**، و محل چندک ما دقیقاً توسط فراوانی تجمعی نسبی وجود داشته باشد، تو این حالت:* چندک مورد نظر ما =.....	۲۹۷ ۲۹۹
روز۲۰	**حل فیشهای : ۲۸۹-۲۹۴-۲۹۷-۳۰۰**	
روز ۲۱	رابطه بین انواع میانگین: میانگین ..... $\geq$ میانگین ..... $\geq$ میانگین ..... اگر ....., هر سه میانگین فوق با هم برابر میشن.	۳۰۴

حالا خود فلش‌کارت‌ها:



« **شاخص‌ها (یا معیارهای عددی)** »

از مشخص‌کننده‌های عددی (شاخص‌های عددی) به چه منظوری استفاده می‌شود؟

آمار کاربردی ۱، عالِم تبریز و احمد هژبر، ص ۱۵

آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار، ص ۵۷

مقدمه:

همان طور که گفته‌ایم، طبقه‌بندی، توصیف و تحلیل داده‌ها در آمار توصیفی با سه روش زیر انجام می‌گیرد:

۱) تهیه جداول توزیع فراوانی: که تو فصل قبل آنها را بررسی کردیم.

۲) توصیف هندسی مشاهدات (رسم نمودارهای آماری): که آنها را در فصلهای بعد مورد بررسی قرار می‌دهیم.

۳) محاسبه شاخص‌های آماری: که در ادامهٔ این فصل آنها را بررسی خواهیم نمود.

پاسخ سؤال) اگر بخواهیم دو جامعه نوعی را با هم **مقایسه** کنیم، این مقایسه را می‌تونیم از طریق روشهای ۱ و ۲ در بالا انجام بدیم (یعنی مقایسه **جداول** فراوانی این دو جامعه و یا مقایسه **نمودار** مشاهدات این دو جامعه)؛ ولی این مقایسهٔ ما حالت کیفی (غیرعددی) داره.

*بنابراین برای اینکه بتوانیم دو جامعهٔ نوعی را به صورت کمّی (عددی) با هم مقایسه کنیم، باید از **اعدادی** استفاده کنیم که به این اعداد،* **شاخص‌ها (مشخص کننده‌ها)ی عددی** می‌گوئیم. پس:

شاخص‌ها (مشخص کننده‌ها)ی عددی، اعدادی هستند که به منظور **مقایسه** کمّی بین دو یا چند جامعه بکار می‌روند.

«انواع شاخص‌ها (یا معیارهای عددی) در علم آمار»

انواع شاخص‌های عددی کدامند؟

هر یک از آنها چه چیزی را نشان می‌دهند؟

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۱

انواع شاخص‌های عددی عبارتند از:

۱- معیارهای ***مرکزی*** ۲- معیارهای ***پراکندگی مطلق***
۳- معیارهای ***پراکندگی نسبی***

که در ادامه به توضیح مختصر هر یک می‌پردازیم.

۱) معیارهای مرکزی (مکانی): شاخص‌های مرکزیت توزیع:

این شاخص‌ها میزان ***تمرکز (مرکزیت)*** مشاهدات را حول یک نقطه نشان می‌دهند، یعنی مشخص می‌کنند که بیشتر مشاهدات در کجا **متمرکز** شده‌اند. *از جمله شاخص‌های مرکزی عبارتند از: میانگین، میانه، مد و ...*

۲) معیارهای پراکندگی مطلق (شاخص‌های تغییرپذیری):

این شاخص‌ها میزان ***پراکندگی*** مشاهدات را نسبت به یک مبداء دلخواه (مانند میانگین: μ) نشان می‌دهند، یعنی مشخص می‌کنند که مشاهدات، نسبت به یک مبداء دلخواه، تا چه اندازه **پراکنده** هستند.

از جمله شاخص‌های پراکندگی می‌توان به دامنه تغییرات (R)، انحراف معیار (δ) و واریانس (δ^2) اشاره کرد.

۳) معیارهای پراکندگی نسبی (شاخص‌های تعیین شکل منحنی توزیع):

از این شاخص‌ها برای مقایسه پراکندگی دو جامعه و یا مقایسه شکل منحنی توزیع دو جامعه به لحاظ تقارن یا چولگی و به لحاظ کشیدگی استفاده می‌شود.

(ادامه تو فیش بعد)

از جمله شاخص‌های پراکندگی نسبی می‌توان به ضریب تغییرات (CV)، ضریب چولگی (SK) و ضریب کشیدگی (E) را نام برد.

توجه: در ادامه این فصل، تنها معیارهای **مرکزی** را مطرح خواهیم کرد و ۲ معیار دیگر را در فصلای بعد بررسی می‌کنیم.

انواع شاخص‌های عددی

- ۱- شاخص‌های **مرکزی**
  - ۱- مُد (نما)
  - ۲- میانه
  - ۳- چندک‌ها
    - ۱- چارک
    - ۲- دهک
    - ۳- صدک
  - ۴- میانگین
- ۲- شاخص‌های پراکندگی **مطلق**
  - ۱- دامنه تغییرات (R)
  - ۲- انحراف معیار (δ)
  - ۳- واریانس (δ^2)
  - ...
- ۳- شاخص‌های پراکندگی **نسبی**
  - ۱- ضریب تغییرات (ضریب پراکندگی): CV
  - ۲- ضریب چولگی (Sk)
  - ۳- ضریب کشیدگی (E)

ایستگاه تغذیه:

بادام زمینی:

بادام زمینی حاوی ماده‌ای مغذی به نام رسوراترول است.

این ماده، آنتی‌اکسیدانی طبیعی است که گیاه بادام زمینی آن را تولید می‌کند.

این ماده دارای خواص ضد پیری است و برای سلامتی ما مفید است.

«شاخص‌ها (یا معیارهای) مرکزی»

۱) *تعریف معیار (یا شاخص)* **مرکزی**:

۲) *انواع شاخص‌های* **مرکزی (مکانی):**

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲۱

آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار،ص ۵۹

«معیار مرکزی: Measure of Central»

۱) تعریف: هر شاخص یا معیار عددی که نشانگر **مرکز داده‌های جامعه** باشه و مشاهدات جامعه، در **اطرافش** قرار بگیرن، یه معیار **مرکزی** است.

به عبارت دیگر هدف از محاسبه شاخص‌های مرکزی این است که ببینیم اکثر داده‌های جامعه در اطراف کدوم داده متمرکز شده‌اند، یعنی به سمت چه عددی گرایش دارند.

---

۲) انواع معیارها یا شاخص‌های مرکزی (مکانی = گرایش به مرکز):

۱- مُد (Mode): *ضعیف‌ترین و کم اهمیت‌ترین شاخص مرکزی است.*

۲- میانه (Median)

۳- چندک‌ها (Quantile): شامل سه دسته چارک‌ها، دهک‌ها و صدک‌ها هستند.

۴- میانگین (Mean): **مهمترین** شاخص مرکزی است و شامل ۳ نوعه:

۱- میانگین **حسابی** ۲- میانگین **هارمونیک** ۳- میانگین **هندسی**

(ادامه تو فیش بعد)

(صرفا برای یادگیری بیشتر):

ویژگی‌های یک شاخص مرکزی خوب و مناسب:

۱- حساب کردنش ساده باشد.

۲- به سادگی قابل درک باشد.

۳- در محاسبه آن از تمامی داده‌ها استفاده شود (مثل شاخص میانگین).

٤- بتونیم بر روی آن عملیات جبری (+، -، × و ÷ و ...) را انجام بدیم (باز هم مثل: میانگین)

٥- مقداری که برای این شاخص حساب می‌کنیم از یک نمونه به نمونهٔ دیگر تغییرات زیادی نکنه، یعنی یه شاخص خوب، شاخصیه که **تغییرپذیری و پراکندگیش کم باشه** (باز هم مثل میانگین).

ایستگاه تغذیه:

بادام زمینی:

آهن موجود در بادام زمینی، برای عملکرد گلبول‌های قرمز خون ما مفید است و

کلسیم موجود در آن هم، به داشتن استخوان‌هایی سالم کمک می کند.

(مهم)

تعریف شاخص مرکزی **مُد (نُما)؛ Mode**

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲۲

آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار،ص ۹۰

۱) مـــد = نمـــا (Mode): داده‌ای (xای) اســـت کـــه **بیشـــترین فراوانی (تکـرار)** را نسـبت بـه سـایر داده‌هـا داشـته باشـد. مـد را بـه اختصار با نماد Mo نمایش می‌دهیم.

نکته ۱) مُد هر دسته از مشاهدات، اندازه‌ای از صفت متغیر (x_i) است؛ یعنی مُد (در صورت وجود)، **حتماً یکی از مشاهدات** است.

مثلاً اگر مشاهدات ۵، ۴، ۲ و ۲ و ۱ را داشته باشیم، آنگاه: $Mo = ۲$

زیرا این مشاهده $(x_i = ۲)$ **بیشترین تکرار (فراوانی)** را دارد (۲ بار تکرار شده).

نکته مهم: همان طور که در این مثال مشاهده کردیم مُد، یکی از داده‌های ما است (یعنی یکی از مقادیر متغیر x است) یعنی $Mo = ۲$ در بین مشاهدات ما وجود دارد و **هرگز امکان نداره که مُد، عددی باشه که در بین داده‌ها نیست** (مثل عدد ۳ و ۶ که در بین داده‌های بالا نیست).

اما در مورد سایر شاخص‌های مرکزی مانند میانه و میانگین بعداً خواهیم دید که گاهی اوقات میانه و میانگین ممکن است در بین خود مشاهدات وجود نداشته باشند، ولی مُد (در صورت وجود) حتماً یکی از مشاهدات است.

نکته ۲) مُد (Mode)، کلمه‌ای فرانسوی است که به معنای **رایج‌ترین** لباس یا سبک می‌باشد.

«محاسبهٔ مد در داده‌های نوع اول (حجم کم)»

(مدیریت ۷۴)

نمای اعداد ۹، ۱۴، ۱۰، ۹، ۵، ۳ و ۲ کدام عدد می‌باشد؟

۱) ۵

۲) ۷

۳) ۹

۴) ۱۴

گزینه ۳)

مُد را می‌تونیم برای هر ۳ نوع داده‌ها حساب کنیم، یعنی هم برای داده‌های نوع **اول** (حجم کم)، هم برای نوع **دوم** (حجم زیاد و تنوع کم) و هم برای نوع **سوم** (حجم زیاد و تنوع زیاد).

در ادامهٔ این فیش، ما مثالی را برای داده‌های نوع **اول** ذکـــر می‌کنیم و در فیش‌های بعدی، نحوهٔ محاسبه مُد برای داده‌های نوع **دوم و سوم** را نشان می‌دهیم.

برای شمارش فراوانی (تکرار) داده‌های خام (داده‌های نوع اول) برای پرهیز از اشتباه در شمارش، بهتر است ابتدا داده‌ها را ***از کوچک به بزرگ*** کنار هم قرار دهیم:

x_i : ۲,۳,۵,۹,۹,۱۰,۱۴ مرتب شده

۹ = مد

با توجه به داده‌های مرتب شده فوق، در این سؤال مد داده‌ها برابر ۹ است، زیرا همه داده‌ها یک بار تکرار شده‌اند، اما عدد ۹ بیشتر از بقیه (۲ بار) تکرار شده است.

توجه: همون‌طور که می‌بینین مُد (در صورت وجود)، ***حتماً یکی از مشاهدات است.***

محاسبهٔ مد برای داده‌های نوع دوم (حجم زیاد، تنوع کم):

مد (نما) مربوط به مشاهدات جدول زیر کدام است؟

x_i	-۱	۰	۱	۲	
F_i	۳۰	۴۰	۲۰	۱۰	$N=\sum F_i=۱۰۰$
f_i	۰/۳	۰/۴	۰/۲	۰/۱	$\sum f_i=۱$

۱) ۱- ۲) ۰ ۳) ۱ ۴) ۲

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲۲

محاسبهٔ مد برای داده‌های نوع دوم (حجم زیاد، تنوع کم):

x_i	-۱	۰	۱	۲	
F_i	۳۰	۴۰	۲۰	۱۰	$N=\sum F_i=۱۰۰$
f_i	۰/۳	۰/۴	۰/۲	۰/۱	$\sum f_i=۱$

پاسخ: گزینه ۲ صحیح است: $Mo=۰$

زیرا در این جدول، دادهٔ $x_i=۰$ بیشترین فراوانی مطلق ($F_i=۴۰$) را دارد و در نتیجه بیشترین فراوانی نسبی (f_i) را در بین داده‌ها دارد.

توجه: در این مثال نیز $Mo=۰$، در بین مشاهدات وجود دارد.

نتیجهٔ مهم: از این مثال نتیجه می‌گیریم که داده‌ای که بیشترین تکرار را داشته باشه، مطمئناً بیشترین فراوانی **مطلق** (F_i) و در نتیجه بیشترین فراوانی **نسبی** (f_i) را خواهد داشت.

«محاسبه مد در مقیاس‌های مختلف»

۱) *شاخص مرکزی مد (نما) برای کدام یک از مقیاسها قابل محاسبه است؟*

۱) اسمی ۲) اسمی و ترتیبی
۳) فاصله‌ای و نسبی ۴) اسمی، ترتیبی، فاصله‌ای و نسبی

---

۲) *برای متغیرهایی با مقیاس* **کیفی اسمی**، *کدام یک از شاخص‌های مرکزی را می‌توان محاسبه نمود؟*

۱) مد و میانه ۲) میانه و میانگین
۳) فقط مد را می‌توان محاسبه کرد. ۴) هیچکدام

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۲

«محاسبه مد در مقیاس‌های مختلف»

دو مثالی که در فیشـــهای قبل حل کردیم، محاســـبه مد را برای داده‌های **کمّی** نشـــان می‌دادند، ولی باید توجه داشـــت که مد را می‌تونیم برای داده‌هایی با مقیاس **کیفی** (*اسمی و ترتیبی*) هم حساب کرد.

مثال زیر محاسبه مد را برای یک متغیر **کیفی اسمی** (گروه خونی) نشان می‌دهد:

مثال ۱(مقیاس کیفی اسمی): *توزیع فراوانی گروه خونی ۳۵ نفر بصورت زیر است. مد یا نما را مشخص کنید:*

گروه خونی	A	B	O	AB	
فراوانی	۵	۷	۱۳	۱۰	$N = ۳۵$

۱) A ۲) B ۳) O ۴) AB

پاسخ: گزینه ۳ صحیح است: ($Mo = O$) ، زیرا گروه خونی «O» **بیشترین فراوانی** را دارد.

اکنون به ذکر مثالی برای محاسبه مد برای یک متغیر **کیفی ترتیبی** (مثل مدرک تحصیلی) می‌پردازیم:

(ادامه تو فیش بعد)

مثال۲(مقیاس کیفی ترتیبی): مد یا نما را برای جدول فراوانی زیر مشخص کنید:

مدرک تحصیلی	دیپلم	لیسانس	فوق لیسانس	دکترا	
فراوانی نسبی	۰/۲	۰/۴	۰/۳	۰/۱	$\sum f_i = ۱$

۱) دیپلم ۲) لیسانس ۳) فوق لیسانس ۴) لیسانس و فوق لیسانس

پاسخ: ***گزینه ۲:*** (لیسانس = Mo)، زیرا افراد دارای مدرک لیسانس ***بیشترین فراوانی نسبی*** (f_i) را دارند.

نتیجه مهم: شاخص مرکزی مد (نما) را می‌توان برای **همهٔ** مقیاس‌ها (***اسمی، ترتیبی، فاصله‌ای و نسبی***) محاسبه نمود (پاسخ سؤال ۱).

نکته: **پاسخ سؤال ۲)**

مُد (نما) تنها شاخص مرکزی است که می‌توان آن را برای داده‌هایی با مقیاس کیفی اسمی حساب کرد.

زیرا همان طور که بعداً خواهیم گفت برای متغیرهای کیفی اسمی ***نمی‌توان*** سایر شاخصهای مرکزی مثل میانه و میانگین را محاسبه نمود.

دانستنی‌های مفید:

برای جلوگیری از یخ زدگی پا در حین انجام فعالیت ورزشی در هوای سرد، از پوشیدن کفش‌های تنگ بپرهیزید.

زیرا کفش تنگ، یکی از عوامل اصلی یخ زدگی پاها است.

(مهم)

با توجه به جدول توزیع فراوانی زیر، مشخص کنید مُد این ۱۰۰ داده کدام است؟

x_i	-۱	۰	۱	۲
F_i	۱۰	۵۰	۹۰	۱۰۰

۱) ۲

۲) ۱۰۰

۳) ۰ و ۱

۴) ۱ و ۲

تست تألیفی

گزینه ۳) با توجه به اینکه فراوانی آخرین طبقه $(x_i = ۲)$ برابر $N = ۱۰۰$ است، پس می‌توان نتیجه گرفت که این جدول، **فراوانی تجمعی** را نشان می‌دهد و نه فراوانی مطلق یا نسبی را.

بنابراین چون مد، داده‌ای است که بیشترین فراوانی مطلق یا نسبی را دارد، پس برای تعیین مد، اول باید فراوانی **مطلق** داده‌ها را بدست آوریم.

نکته: فراوانی مطلق هر طبقه از تفاضل فراوانی تجمعی آن طبقه از فراوانی تجمعی طبقه قبل بدست می‌آید:

$$F_i = F_{c_i} - F_{c_i - ۱}$$

x_i	-۱	۰	۱	۲
فراوانی تجمعی F_{c_i}	۱۰	۵۰	۹۰	N=۱۰۰
فراوانی مطلق F_i	۱۰	50-10=40	۹۰-50=40	100-90=10

بنابراین مشاهده می‌شود که ۱ و ۰=Mo است، یعنی جامعه ما دو مدی است، زیرا مشاهدات ۰ و ۱ بیشتر از سایر مشاهدات تکرار شده‌اند(۴۰ بار).

نکته مهم: باید توجه داشت که مد، **همیشه از بین x_iها** انتخاب می‌شود **و نه از بین F_iها یا f_iها**. یعنی مد، یک مشاهده (x_i) است نه یک فراوانی (F_i یا f_i).

(برنامه‌ریزی شهری (۸)

برای داده‌های زیر:

x	۲	۳	۴	۵	۶
f	۳	۴	۸	۷	۵

مقدار نما کدام است؟

۱) ۳

۲) ۴

۳) ۵

۴) ۸

گزینه ۲) زیرا $x_i = ۴$ دارای بیشترین فراوانی است $(f_i = ۸)$

x	۲	۳	۴	۵	۶
f	۳	۴	۸	۷	۵

توجه: منظور طراح سؤال از f، در واقع همان فراوانی **مطلق** (F) است. زیرا همان طور که قبلاً گفته‌ایم، در بعضی کتاب‌ها، فراوانی مطلق را با f و فراوانی تجمعی را با F نمایش می‌دهند، که البته تشخیص آنها به سادگی از روی اعداد جدول امکان‌پذیر است، زیرا همان طور که می‌دانیم:

۱) فراوانی **مطلق** همواره عددی ***صحیح و بزرگتر مساوی یک*** است:

$F_i \geq ۱$

۲) فراوانی **نسبی** همواره عددی اعشاری (کسری) و ***بین صفر و یک*** است:

$۰ < f_i < ۱$

۳) فراوانی **تجمعی** نیز همواره حالت ***صعودی (افزایشی)*** دارد، یعنی اعداد جدول، از چپ به راست افزایش می‌یابند تا جایی که فراوانی تجمعی طبقه آخر برابر N = تعداد داده‌ها می‌شود.

نکته: گزینه ۴ غلط است، زیرا یک فراوانی رو نشون میده (فراوانیِ x=۴) و نه یه x رو، در حالیکه می‌دانیم مد، **یک داده** (یعنی x_i) است که بیشترین فراوانی را دارد و **نه یک فراوانی** (F_i یا f_i).

(مهم):

ویژگی‌های مُد (نما):

ابتدا شاخص مرکزی مُد را برای مثال زیر بدست آورید.

ثانیاً بیان کنید که این مثال، کدام ویژگی شاخص مد را نشان می‌دهد؟

مثال: *مد (نما) مربوط به مشاهدات*

۷ و ۲ و ۹ و ۲ و ۷ و ۲ و ۴ و ۳ و ۷ و ۳ کدام است؟

۱) ۲ ۲) ۷ ۳) ۲ و ۷ ۴) وجود ندارد.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲۲ و ۲۳

آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار،ص ۹۱

پاسخ مثال) گزینه «۳»

نکته ۱: شاخص مرکزی مُد (نما) ***لزوماً منحصر بفرد نیست***، یعنی ممکن است بیش از ۱ مُد داشته باشیم.
یعنی اگر دو یا چند داده (**نه همه داده‌ها**) بیشترین تکرار را داشته باشند، اون‌وقت **همه** آن داده‌ها را به عنوان مد مشاهدات در نظر می‌گیریم.

در این مثال، مشاهداتِ $x_i = ۲,۷$، بیشترین تکـرار را نسبت به بقیه داده‌ها دارند (۳ بار)، پس این توزیع **دو مُد** دارد:

$$x_i : ۳,۷,۳,۴,۲,۷,۲,۹,۲,۷ \Rightarrow \boxed{Mo = ۲,۷}$$

توجه: اگر دو یا چند مد داشته باشیم توزیع را **دو مُدی (دو نمایی) یا چند مُدی (چند نمایی)** می‌گوییم و در این صورت شاخص مد، دیگر نقطه تمرکز داده‌ها نیست.

(مهم):

ویژگی‌های مُد (نما):

ابتدا شاخص مرکزی مُد را برای مثال زیر بدست آورید.

ثانیاً بیان کنید که این مثال، کدام ویژگی شاخص مد را نشان می‌دهد؟

مثال : *مُد (نما) مربوط به مشاهدات*

۷ و ۲ و ۲ و ۴ و ۷ و ۲ و ۴ و ۷ و ۴ کدام است؟

۱) ۷ ۲) ۳

۳) ۴ ۴) وجود ندارد

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۲ و ۲۳

پاسخ مثال) گزینه ۴

نکته: اگر **تمامی** داده‌ها (x_i ها) به یک اندازه تکرار شده باشند (یعنی **فراوانی مطلق یا نسبی یکسانی** داشته باشند)، آنگاه جامعه **مد (Mo) نخواهد داشت.**

یعنی در این حالت، **مد جامعه تهی** است که آن را به صورت زیر نشان می‌دهیم: $Mo = \{\ \}$ یا $Mo = \{\phi\}$

در این مثال، تمام، داده‌ها (۲، ٤ و ۷) ***به یک اندازه تکرار شده‌اند*** *(۳ بار)، یعنی* ***فراوانی مطلق و نسبی یکسانی*** *دارند، بنابراین این مشاهدات* ***مد ندارند*** *(مد این مشاهدات تهی است):* $Mo = \{\ \}$ یا $Mo = \{\phi\}$

$x_i = ۴,۷,۴,۲,۷,۴,۲,۲,۷$

(مهم):

ویژگی‌های مُد (نما):

ابتدا شاخص مرکزی مُد را برای مثال زیر بدست آورید.

ثانیاً بیان کنید که این مثال، کدام ویژگی شاخص مد را نشان می‌دهد؟

مثال: *اگر مشاهدات ما ٦ و ٥ و ٥ و ٢ و ١ باشند آنگاه مد مشاهدات* $(-\frac{۱}{۲}x_i + ۳)$ *کدام است؟*

۱) ۵ ۲) $-\frac{۵}{۲}$ ۳) ۸ ۴) $\frac{۱}{۲}$

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۲ و ۲۳

گزینه ۳) با توجه به اینکه فراوانی آخرین طبقه $(x_i = ۲)$ برابر $N = ۱۰۰$ است، پس می‌توان نتیجه گرفت که این جدول، **فراوانی تجمعی** را نشان می‌دهد و نه فراوانی مطلق یا نسبی را.

بنابراین چون مد، داده‌ای است که بیشترین فراوانی مطلق یا نسبی را دارد، پس برای تعیین مد، اول باید فراوانی **مطلق** داده‌ها را بدست آوریم.

نکته: فراوانی مطلق هر طبقه از تفاضل فراوانی تجمعی آن طبقه از فراوانی تجمعی طبقه قبل بدست می‌آید:

$$F_i = F_{c_i} - F_{c_{i-۱}}$$

x_i	-۱	۰	۱	۲
فراوانی تجمعی F_{c_i}	۱۰	۵۰	۹۰	N=۱۰۰
فراوانی مطلق F_i	۱۰	50-10=40	۹۰-50=40	100-90=10

بنابراین مشاهده می‌شود که ۱ و ۰=Mo است، یعنی جامعه ما دو مدی است، زیرا مشاهدات ۰ و ۱ بیشتر از سایر مشاهدات تکرار شده‌اند(۴۰ بار).

نکته مهم: باید توجه داشت که مد، **همیشه از بین x_iها** انتخاب می‌شود **و نه از بین F_iها یا f_iها**. یعنی مد، یک مشاهده (x_i) است نه یک فراوانی (F_i یا f_i).

(مهم):

اگر کل داده‌ها بر ۲- تقسیم شوند، مد جامعه چه تغییری خواهد کرد؟

۱) ۲برابر می‌شود.

۲) تغییری نمی‌کند.

۳) نصف می‌شود.

۴) $-\frac{۱}{۲}$ برابر می‌شود.

توجه: پس از پاسخگویی به این سؤال، حتماً به نکته‌ای که در پاسخ سؤال آمده ، دقت کنید.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۳

گزینه ۴) اگر کل داده‌ها بر عددی تقسیم شوند، مد این داده‌ها هم بر همین عدد تقسیم می‌شود:

$$\begin{cases} x_i : x_1, x_2, \ldots, x_n \Rightarrow y_i = -\frac{1}{2} x_i \\ \Downarrow \\ Mo = Mo_{x_i} \qquad \boxed{Mo_y = -\frac{1}{2} Mo_x} \end{cases}$$

توجه: اگر در صورت سؤال گفته می‌شد، تنها **برخی** از داده‌ها بر ۲- تقسیم شوند، آنگاه دیگر **نمی‌توانستیم** با قاطعیت بگوئیم که مد بر ۲- تقسیم می‌شود.

مثلاً اگر مشاهدات ما به صورت ۷ و ۶ و ۶ و ۴ و ۲ و ۱ باشند، آنگاه:

۱) اگر **تنها دو دادهٔ** $x_i = ۲,۴$، نصف شوند، مد مشاهدات تغییری نخواهد کرد و هم چنان $Mo = 6$ خواهد بود.

۲) اما اگر **فقط مُد** (یعنی دو دادهٔ $x_i = ۶,۶$) را نصف کنیم، آنگاه اگرچه تمام مشاهدات تغییر نکرده‌اند، ولی چون مُد تغییر کرده، پس مد مشاهدات جدید (یعنی مُد داده‌های: ۷ و ۳ و ۳ و ۴ و ۲ و ۱) هم تغییر می‌کند، یعنی در این حالت مد نیز نصف می‌شود:

$$Mo_{y_i} = \frac{1}{2} Mo_x \Rightarrow \frac{1}{2}(6) = \frac{6}{2} = 3$$

نتیجه: تغییر مد، به این امر بستگی دارد که کدام داده تغییر کرده.

مهم:

خاصیت مهم مُد (نما):

*از آنجا که مد، مشاهده‌ای است با بیشترین فراوانی، پس هنگامی که از مد به عنوان نماینده و تخمین مقدار عددی تک تک مشاهدات آماری استفاده می‌کنیم، آنگاه ..... **حداقل** می‌شود.*

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲۲ و ۲۳

جواب: تعداد تخمین‌های اشتباه (تعداد خطاها در تخمین : N_e)

توضیح: از آنجا که مُد مشاهده‌ای است که بیشترین فراوانی را دارد، پس اگر از **مد** به عنوان نماینده (برای تخمین مقدار عددی تک تک مشاهدات) استفاده کنیم، آنگاه ***تعداد تخمین‌های اشتباه (تعداد خطاها در تخمین: N_e) حداقل خواهد بود.***

به عبارت دیگر اگر تعداد تخمین‌های اشتباه را با N_e نشان دهیم، به ازای استفاده از **مد بجای تک تک داده‌ها،** $N_e = \min$ خواهد شد.

مثال: داده‌های مقابل را در نظر بگیرید: ۵ و ۹ و ۹ و ۹ و ۹ و ۲ و ۲

در این داده‌ها مد مساوی $Mo = ۹$ است.

۱) اگر بجای تمام داده‌ها مقدار ۲ را قرار دهیم، اون وقت در تخمین ۵ داده (۵ و ۹ و ۹ و ۹ و ۹) اشتباه کرده‌ایم (زیرا تنها دو دادهٔ ۲ و ۲ را درست تخمین زده‌ایم).

۲) و اگر بجای تمام داده‌ها مقدار ۹ (مد مشاهدات) را قرار دهیم، اون وقت فقط در تخمین ۳ داده (۵ و ۲ و ۲) اشتباه می‌کنیم.

۳) و اگر بجای تمام داده‌ها مقدار ۵ را قرار دهیم، اون وقت در تخمین ٦ داده (۹ و ۹ و ۹ و ۹ و ۲ و ۲) اشتباه می‌کنیم.

نتیجه مهم: با قرار دادن مُد (۹) به جای سایر داده‌ها، **تعداد تخمین‌های اشتباه، حداقل می‌شود.**

$$Mo=۹ \longrightarrow N_e = \min$$

تعداد تخمین‌های اشتباه **حداقل**

«محاسبه مد در داده‌های نوع سوم (حجم زیاد و تنوع زیاد)»

مثال ((طبقات پیوسته):

مد در داده‌های جدول زیر کدام است؟

حدود طبقات	۲ -۴	۴ –۶	۶ –۸	۸ –۱۰
فراوانی	۱۰	۱۴	۸	۳

۱) ۵/۵ ۲) ۴/۸ ۳) ۴ ۴) ۵/۲

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲٤

محاسبه مد در داده‌های نوع سوم (حجم زیاد و تنوع زیاد):

در فیشهای قبل با نحـــوهٔ محاســبه مد در داده‌های نوع **اول** (حجـم کم) و نوع **دوم** (حجم زیاد و تنوع کم) آشنا شدیم و اکنون قصد داریم به محاسبه مد در داده‌های نوع **سوم** (حجم زیاد، تنوع زیاد) بپردازیم:

مراحل محاسبه مد در داده‌های نوع سوم:

الف- انتخاب **طبقه مد دار** (طبقه‌ای که بیشترین فراوانی مطلق (F_i) یا فراوانی نسبی (f_i) را دارد).

ب- یافتن حدود **واقعـی** مـد دار (**پیوسـته** کـردن طبقـه مـد دار در صورت گسسته بودن)

ج- محاسبه مد از رابطهٔ زیر:

طول یا فاصله طبقه $\times \dfrac{d_1}{d_1+d_2} +$ **حد پایین واقعی طبقه مد دار = Mo**

$$\Rightarrow \boxed{Mo = L_{Mo} + \frac{d_1}{d_1+d_2} \times I}$$

که در رابطه بالا:

(ادامه تو فیش بعد)

L_{Mo} : حد پایین **واقعی** طبقه مدار

$d_۱$: تفاضل فراوانی طبقه مد دار از فراوانی طبقه **قبل**:

$\boxed{d_۱ = f_i - f_{i-۱}}$ یا $\boxed{d_۱ = F_i - F_{i-۱}}$

$d_۲$: تفاضل فراوانی طبقه مد دار از فراوانی طبقه **بعد**:

$\boxed{d_۲ = f_i - f_{i+۱}}$ یا $\boxed{d_۲ = F_i - F_{i+۱}}$

I: فاصله طبقات که همیشه با طول طبقات برابره.

اکنون با ذکر ۲ مثال، مراحل فوق را گام به گام انجام می‌دهیم.

مثال اول بیانگر زمانی است که طبقات جدول **پیوسته** هستند و بنابراین دیگر **لازم نیست** طبقه مد دار را پیوسته کنیم.

در مثال دوم، طبقات جدول **گسسته** هستند، بنابراین ابتدا بایستی **طبقه مد دار** را پیوسته کنیم:

مثال ۱: (طبقات پیوسته):

مد در داده‌های جدول زیر کدام است؟

حدود طبقات	۲ − ۴	۴ − ۶ (طبقه مد دار)	۶ − ۸	۸ − ۱۰
فراوانی	۱۰ ($d_۱$=۴)	$\boxed{۱۴}$	۸ ($d_۲$=۶)	۳

۱) ۵/۵ ۲) ۴/۸ ۳) ۴ ۴) ۵/۲

(ادامه تو پشت فیش)

پاسخ: گزینه ۲. مراحل کار به صورت زیر است:

الف) طبقه (۶ – ۴) طبقه مد دار است، زیرا **بیشترین فراوانی** را دارد (۱۴). با توجه به طبقه مد دار می‌تونیم بگیم که مد عددی بین ۴ تا ۶ است.

ب) با توجه به اینکه طبقات **پیوسته** هستند، نیازی به پیوسته کردن آنها **نیست**.

ج) محاسبه مد از رابطه زیر:

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I} \Rightarrow Mo = ۴ + \frac{(۱۴-۱۰)}{(۱۴-۱۰)+(۱۴-۸)} \times ۲$$

$$\Rightarrow Mo = ۴ + (\frac{۴}{۴+۶}) \times ۲ \quad \Rightarrow \quad ۴ + (\frac{۴}{۱۰}) \times ۲ \quad \Rightarrow \quad ۴ + \frac{۸}{۱۰}$$

$$\Rightarrow Mo = ۴ + ۰/۸ \quad \Rightarrow \quad \boxed{Mo = ۴/۸}$$

توجه: همواره برای اطمینان از صحت محاسباتمان، باید به این نکته توجه داشته باشیم که مقدار مدی که از فرمول فوق بدست می‌آوریم، **باید حتماً بین حد پایین و بالای طبقه مد دار قرار داشته باشد**.

در مثال فوق نیز مُد (یعنی ۴/۸) بین حد پایین طبقه مد دار (یعنی ۴) و حد بالای این طبقه (یعنی ۶) قرار داره: $\boxed{۴ < Mo = ۴/۸ < ۶}$

«محاسبه مد در داده‌های نوع سوم (حجم زیاد و تنوع زیاد)»

مثال ۲ (طبقات گسسته):

مد داده‌های جدول زیر کدام است؟

حدود طبقات	۲ -۴	۵ –۷	۸ –۱۰	۱۱ –۱۳
فراوانی	۱۰	۱۴	۸	۳

۱) ۵/۲ ۲) ۵/۷ ۳) ۴/۵ ۴) ۶/۵

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲٤

حدود طبقات	۲ — ۴	۵ — ۷ (طبقه مد دار)	۸ — ۱۰	۱۱ — ۱۳
فراوانی	۱۰	$\boxed{۱۴}$	۸	۳

$d_۱=۴$ $d_۲=۶$

۱) ۵/۲ ۲) ۵/۷ ۳) ۴/۵ ۴) ۶/۵

پاسخ: **گزینه ۲**. مراحل محاسبه مد بصورت زیر است:

الف) *طبقه (۷ — ۵) طبقه مد دار است، زیرا* **بیشترین فراوانی** *(۱۴) را دارد.*

ب) *به علت* **گسسته** *بودن طبقات، بایستی حدود واقعی طبقه مد دار را بدست آوریم.*

برای این منظور بایستی نصف فاصله حد بالای طبقه مد دار از حد پایین طبقه بعدیش را بدست آورده و یا نصف فاصله حد پایین طبقه مد دار از حد بالای طبقه قبلیش را بدست بیاریم و سپس این مقدار را به حد بالای طبقه مد دار اضافه و از حد پایینش کم کنیم:

(ادامه تو فیش بعد)

نصف این فاصله $\frac{۱}{۲}$ نصف این فاصله

C - L	۲ — ۴	۵ — ۷	۸ — ۱۰	۱۱ — ۱۳
F_i	۱۰	۱۴	۸	۳

فاصله = ۱ فاصله = ۱

پس بـایـد عـدد $\frac{۱}{۲}=۰/۵$ رو از حـد پـایین طبقـه مـد دار کم کنیم $(۵-۰/۵=۴/۵)$ و بـه حـد بـالای طبقـه مـد دار اضـــافـه کنیم $(۷+۰/۵=۷/۵)$ تا با این کار، اون فاصـــلهٔ یک واحدی که بین طبقات وجود داره، از بین بره و در نتیجه طبقات ما از حالت گسـسـته (جدا از هم) به حالت پیوسته (چسبیده بهم) در بیان، تا بعدش بتونیم مُد رو از طریق فرمولش حساب کنیم.

پس حدود **واقعی** طبقه مد دار برابره با: $\boxed{(۴/۵,۷/۵)}$

در نتیجه جدول **پیوسته** ما بصورت زیر درمیاد:

C - L	۱/۵-۴/۵	۴/۵-۷/۵ (طبقه مد دار)	۷/۵-۱۰/۵	۱۰/۵-۱۳/۵
F_i	۱۰	۱۴	۸	۳

$d_۱=۴$ $d_۲=۶$

(ادامه تو پشت فیش)

ج) محاسبه مقدار مد از رابطه زیر:

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I} \Rightarrow ۴/۵ + \frac{(۱۴-۱۰)}{(۱۴-۱۰)+(۱۴-۸)} \times ۳$$

$$\Rightarrow ۴/۵ + \frac{۴}{۴+۶} \times ۳ \Rightarrow ۴/۵ + \frac{۴}{۱۰} \times ۳ \Rightarrow ۴/۵ + \frac{۱۲}{۱۰} \Rightarrow$$

$$۴/۵ + ۱/۲ \Rightarrow \boxed{Mo = ۵/۷}$$

توجه: طول طبقات که با **فاصله** طبقات برابره، بصورت زیر بدست میاد:

$\Longrightarrow$ تفاضل حد بالای دو طبقه متوالی = تفاضل حد پایین دو طبقه متوالی= I

فاصله (طول) طبقات:

$$I = L_۲ - L_۱ = U_۲ - U_۱ \Rightarrow$$

$$I = ۵ - ۲ = ۷ - ۴ \Rightarrow \boxed{I = ۳}$$

چند نکتهٔ خفن، ظریف و باحال دربارهٔ مد:

نکته ۱) قبلاً در مورد محاسبه مد برای داده‌های نوع اول و دوم، گفتیم که مد حتماً در بین داده ها قرار دارد، اما برای داده های نوع سوم که بصورت **فاصله‌ای** (یعنی بصورت C-L) طبقه‌بندی می شوند، ممکن است مد، اصلاً در بین مشاهدات ما وجود نداشته باشد.

مثلاً در مثال ۱ (در سوال دو- سه فیش قبلی) ممکن است واقعاً دادهٔ $Mo = ۴/۸$ در بین داده‌ها نباشد و یا در مثال ۲ ممکن است لزوماً دادهٔ $Mo = ۵/۷$ در بین داده‌ها نباشد، **(ادامه تو فیش بعد)**

به همین دلیل اســت که در برخی منابع، فرمول مد را بصــورت **تقریبی** (یعنی با نماد $\approx$ به معنای **تقریباً مساوی است با**) نشان می‌دهند:

$$Mo \approx L_{Mo} + \frac{d_1}{d_1 + d_2} \times I$$

نکته ۲) روش معمولی (کم دقت) برای محاسبه مد در داده‌های نوع سوم که بصورت **فاصله‌ای** طبقه‌بندی شده‌اند، این است که **مرکز طبقه مد دار** را به عنوان مد بپذیریم، یعنی مثلاً در مثال ۲، مد از روش معمولی (کم دقت) برابر است با:

$$Mo = \frac{5+7}{2} = \frac{12}{2} = 6$$

اما روش **دقیقتر** استفاده از فرمول زیر است:

$$Mo = L_{Mo} + \frac{d_1}{d_1 + d_2} \times I$$

در واقع این روش سعی شده تا مد توزیع، بجای آنکه دقیقاً در مرکز طبقه مد دار (یعنی در وسط طبقهٔ ۷ – ۵ که برابر ۶ است) واقع شود ، **به سمت کرانه‌ای که همسایه آن فراوانی بیشتری دارد، متمایل‌تر (نزدیک‌تر) باشد.**

در مثال ۲ نیز همسایه کران ۷ دارای فراوانی ۸ و همسایه کران ۵ دارای فراوانی ۱۰ است، بنابراین مد، به کرانه ۴ که فراوانی بیشتری دارد، متمایل‌تر و نزدیک‌تر است (تا به کرانه ۸ که فراوانی کمتری دارد). **(ادامه تو پشت فیش)**

اما اینکه این تمایل به چه اندازه باشد به نسبت $d_۱$ و $d_۲$ در فرمول مد بستگی دارد.

نکته ۳) اگر **طبقه اول** جدول، طبقه مد دار باشد، آنگاه:

فراوانی طبقه مد دار $= F_۱ = d_۱ =$ تفاضل فراوانی طبقه مد دار از فراوانی طبقه **قبل**

زیرا در این حالت، قبل از طبقه اول، دیگر طبقه‌ای وجود ندارد که فراوانی‌ای داشته باشد، بنابراین می‌تونیم فراوانی طبقه ماقبل اول رو صفر در نظر بگیریم و در نتیجه بنویسیم:

$$d_۱ = F_{\text{اول}} - F_{\text{ماقبل اول}} \Rightarrow d_۱ = F_۱ - ۰ \Rightarrow \boxed{d_۱ = F_۱}$$

برای درک بهتر این موضوع به مثال زیر توجه کنین:

مثال (برای حالتی که طبقه مد دار، طبقه اول باشه):

C –L	طبقه مد دار ۲ – ۴	۴ – ۶	۶ – ۸
F_i	۱۰	۵	۴

$d_۲=۵$

$d_۱=۱۰$

پس مد برابره با:

(ادامه تو فیش بعد)

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I} \Rightarrow Mo = ۲ + \frac{۱۰}{۱۰+۵} \times ۲ \Rightarrow$$

$$۲ + \frac{۱۰ \times ۲}{۱۵} \Rightarrow Mo = ۲ + \frac{۲۰}{۱۵} \Rightarrow ۲ + \frac{۴}{۳} \Rightarrow ۲ + ۱/۳۳$$

$$\Rightarrow \boxed{Mo = ۳/۳۳}$$

توجه کنین که: در مثال بالا، چون **طبقه اول**، طبقه مد دار بود، پس $d_۱$ با $F_۱$ مساوی شد، یعنی:

$$d_۱ = F_۱ - F_{\text{ماقبل اول}} \Rightarrow d_۱ = ۱۰ - ۰ = ۱۰ \Rightarrow \boxed{d_۱ = F_۱ = ۱۰}$$

و اگر **طبقه آخر**، طبقه مد دار باشد، آنگاه:

فراوانی طبقه مد دار : $F_{\text{آخر}} = d_۲ =$ تفاضل فراوانی طبقه مد دار از فراوانی طبقه **بعد**

زیرا در این وضعیت، بعد از طبقه آخر، دیگر طبقه‌ای وجود ندارد، در واقع فراوانی طبقات بعدی صفر است:

$$d_۲ = F_{\text{آخر}} - F_{\text{بعدی}} \Rightarrow d_۲ = F_{\text{آخر}} - ۰ \Rightarrow d_۲ = F_{\text{آخر}}$$

(ادامه تو پشت فیش)

مثال (برای حالتی که طبقه مد دار، طبقه آخر باشه):

C - L	2 − 4	4 − 6	$\overbrace{6-8}^{\text{طبقه مد دار}}$
F_i	۴	۵	۱۰

$d_۱=۵$ $d_۲=۱۰$

$$Mo = ۶ + \frac{۵}{۵+۱۰} \times ۲ \Rightarrow ۶ + \frac{۵ \times ۲}{۱۵} \Rightarrow ۶ + \frac{۱۰}{۱۵}$$

$$\Rightarrow Mo = ۶ + \frac{۲}{۳} \Rightarrow ۶ + ۰/۶۶ \Rightarrow \boxed{Mo = ۶/۶۶}$$

توجه کنین که: تو مثال بالا $d_۲$ با $F_{\text{آخر}}$ برابر شد:

$$d_۲ = F_{\text{آخر}} - F_{\text{بعدی}} \Rightarrow d_۲ = ۱۰ - ۰ = ۱۰ \Rightarrow \boxed{d_۲ = F_{\text{آخر}} = 10}$$

(مهم):

(مدیریت ۷۶)

داده‌های طبقه‌بندی شده زیر را در نظر بگیرید. مقدار مد را محاسبه کنید.

فاصله طبقات CL	-۱۰-۰	۰-۱۰	۱۰-۲۰
فراوانی F_i	۳۰	۲۰	۲۰

۱) -۱۰

۲) ۲/۵

۳) -۲/۵

۴) ۰

گزینه ۳)

گام ۱) مشخص کردن طبقه مد دار:

C-L	۰ — ۱۰ (طبقه مد دار)	۱۰ — ۲۰ ($I=۱۰$)	۲۰ — ۳۰
F_i	$d_۱=۳۰$	$d_۲=۱۰$	۲۰

یادآوری: طبقه مـد دار، طبقه اوّله، پس $d_۱$ را مسـاوی $F_۱$ در نظـر می‌گیریم (یعنی: $d_۱ = F_۱ = ۳۰$)، چون فراوانی طبقه ماقبل اول، صفر در نظر گرفته میشه: $d_۱ = F_۱ - ۰ \Rightarrow d_۱ = ۳۰ - ۰ = ۳۰$

گام ۲) پیوسته کردن طبقات (در صورت نیاز): طبقات ما پیوسته هستن و حدود واقعی طبقه مد دار به صورت ۱۰,۰- است.

همچنین چون طبقات ما پیوسته هستن، پس فاصله طبقات (I) با عرض طبقات برابره، پس: $\boxed{I = ۱۰ - ۰ = ۱۰}$

گام ۳) محاسبه مُد:

$$\text{Mo} = \text{حد پایین } \textbf{واقعی} \text{ طبقه مد دار} + \frac{d_۱}{d_۱ + d_۲} \times \text{فاصله (طول) طبقه}$$

(ادامه تو فیش بعد)

$$Mo = -۱۰ + \frac{۳۰}{۳۰+۱۰} \times ۱۰ \Rightarrow Mo = -۱۰ + \frac{۳۰}{۴۰} \times ۱۰ \Rightarrow$$

$$Mo = -۱۰ + \frac{۳}{۴} \times ۱۰ \Rightarrow -۱۰ + \frac{۳۰}{۴} \Rightarrow -۱۰ + \frac{۱۵}{۲} \Rightarrow$$

$$Mo = -۱۰ + ۷/۵ \Rightarrow \boxed{Mo = -۲/۵}$$

روش تستی: همون طور که تو فیشای قبل یاد گرفتیم، روش معمولی و کم دقت برای محاسبه مد در جداول طبقه‌بندی شده (که به صورت (L-Uهستن) اینه که **مرکز طبقه مد دار** رو به عنوان مد مشاهدات در نظر بگیریم که تو این سؤال درواقع برابره با مرکز طبقه (۰-۱۰-) که ۵- است:

$$\boxed{\text{مرکز طبقه مد دار} : \frac{-۱۰+۰}{۲} = \frac{-۱۰}{۲} = -۵}$$

اما واقعیت اینه که مقدار واقعی و دقیق مد، دقیقاً در مرکز طبقه مد دار قرار نمی‌گیره، بلکه **به سمت طبقهٔ مجاوری که بیشترین فراوانی رو داره متمایل میشه.**

یعنی تو این سؤال، طبقات مجاور طبقه اول، یکی طبقه ماقبل اوله که مسلماً فراوانیش صفره و دیگری، طبقه دومه که فراوانیش ۲۰ است. پس مقدار غیردقیق مد که ۵- بود، باید به سمت طبقه دوم که فراوانی بیشتری نسبت به طبقه ماقبل اول داره، نزدیک بشه **(ادامه تو پشت فیش)**

و این یعنی اینکه مقدار مد باید کمی از ۵- بزرگتر باشه که چنین وضعیتی فقط تو گزینه ۳ $(Mo = -۲/۵)$ دیده میشه.

نکته تستی : گزینه‌های ۲ و ۴ از همون اول و بدون هیچ محاسبه‌ای قابل حذف بودن، چون ما می‌دونیم که مقدار مد، همیشه باید بین حد پایین و بالای طبقه مد دار باشه،

یعنی تو این سؤال، مد باید حتماً مقدارش بین ۱۰- و ۰ باشه $(-۱۰ < Mo < ۰)$ ولی اعداد ۲/۵ و ۰ در این فاصله قرار ندارن.

(GI S ۸۵)

در جدول داده‌های زیر، مُد جامعه کدام است؟

x_i	۱۲-۱۶	۱۶-۲۰	۲۰-۲۴	۲۴-۲۸	۲۸-۳۲
f_i	۵	۱۲	۱۸	۹	۶

۱) ۲۱/۳

۲) ۲۱/۶

۳) ۲۲/۱

۴) ۲۲/۴

گام ۱) مشخص کردن طبقه مد دار : طبقه سوم (۲۴-۲۰)، طبقه مد داره، چون **بیشترین فراوانی** رو داره: $F_i = ۱۸$

C-L	I=۴ ۱۲-۱۶	۱۶-۲۰	طبقه مد دار ۲۰-۲۴	۲۴-۲۸	۲۸-۳۲
F_i	۵	۱۲ $\xleftrightarrow{d_۱=۶}$	$\boxed{۱۸}$	$\xleftrightarrow{d_۲=۹}$ ۹	۶

گام ۲) پیوسته کردن طبقات *(در صورت نیاز)*:

طبقات ما پیوسته هستن، پس دیگه نیازی به این کار نیست. بنابراین حدود واقعی طبقه مد دار به صورت (۲۴-۲۰) خواهد بود.

دقت کنین که: چون طبقات ما **پیوسته** هستن، پس فاصله طبقات (I) با عرض طبقات برابره، پس: $\boxed{I = ۱۶ - ۱۲ = ۴}$

گام ۳) محاسبه مد:

فاصله (طول) طبقات $\times \dfrac{d_۱}{d_۱ + d_۲} +$ حد پایین **واقعی** طبقه مد دار = **مُد**

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I}$$

$$\Rightarrow Mo = ۲۰ + \frac{(۱۸ - ۱۲)}{(۱۸ - ۱۲) + (۱۸ - ۹)} \times ۴ \Rightarrow$$

(ادامه تو فیش بعد)

$$\Rightarrow Mo = ۲۰ + \frac{۶}{۶+۹} \times ۴ \Rightarrow ۲۰ + \frac{۶}{۱۵} \times ۴ \Rightarrow$$

$$۲۰ + \frac{۲}{۵} \times ۴ \Rightarrow Mo = ۲۰ + \frac{۸}{۵} \xrightarrow{\frac{۸}{۵} = \frac{۱۶}{۱۰}}$$

$$Mo = ۲۰ + \frac{۱۶}{۱۰} \Rightarrow ۲۰ + ۱/۶ \Rightarrow \boxed{Mo = ۲۱/۶}$$

نکته تستی: مقدار مد از روش معمولی و کم دقت برابر با **مرکز طبقه مد داره**، یعنی:

$$\boxed{\textbf{مرکز طبقه مد دار} : \frac{۲۰+۲۴}{۲} = \frac{۴۴}{۲} = ۲۲}$$

اما مقدار دقیق‌تر مُد به سمت طبقهٔ **مجاوری** که **فراوانی بیشتری** داشته باشه، نزدیک میشه.

در این سؤال، طبقات مجاور طبقه مد دار، طبقات دوم و سوم هستن که طبقه دوم فراوانی بیشتری نسبت به طبقه سوم داره (۹<۱۲)، پس مقدار مد باید از ۲۲ به سمت طبقه دوم متمایل بشه، یعنی باید مقدارش یه خرده کمتر بشه؛ به عبارت بهتر، مقدار دقیق مد باید کمتر از مرکز طبقه مد دار (۲۲) باشه، یعنی $Mo < ۲۲$ *که این وضعیت فقط تو گزینه‌های ۱ و ۲ دیده میشه. بنابراین از همون اول می‌تونستیم گزینه‌های ۳ و ٤ رو خیلی سریع حذف کنیم:*

(ادامه تو پشت فیش)

CL	۱۲-۱۶	طبقه **مجاور** ۱۶-۲۰	طبقه مد دار ۲۰-۲۴	طبقه **مجاور** ۲۴-۲۸	۲۸-۳۲
F_i	۵	$\boxed{۱۲}$	۱۸	۹	۶

مقایسه فراوانی‌ها

۱۲ > ۹

⇓

پس مقدار دقیق مد باید از مرکز طبقه مد دار (۲۲) به سمت **مقادیر کمتر** (یعنی به سمت طبقه دوم) متمایل بشه؛ یعنی مد باید کمتر از ۲۲ باشه.

(حسابداری ۸۴)

در جدول داده‌های زیر اندازه مُد (Mo) کدام است؟

حدود دسته	۸-۱۱	۱۱-۱۴	۱۴-۱۷	۱۷-۲۰	۲۰-۲۳
فراوانی	۴	۱۰	۱۸	۱۲	۶

۱) ۱۵/۵

۲) ۱۵/۶

۳) ۱۵/۷

۴) ۱۵/۸

گام ۱) مشخص کردن طبقه مد دار : طبقه سوم(۱۴-۱۷) ، طبقه مد داره، چون **بیشترین فراوانی** رو داره (۱۸):

C-L	I=۳ ۸-۱۱	۱۱-۱۴	طبقه مد دار ۱۴-۱۷	۱۷-۲۰	۲۰-۲۳
F_i	۴	۱۰ $\xleftrightarrow{d_1=۸}$	$\boxed{۱۸}$	$\xleftrightarrow{d_2=۶}$ ۱۲	۶

گام ۲) پیوسته کردن طبقات (در صورت لزوم):

چون طبقات جدول ما، **پیوسته** هستن، پس دیگه نیازی به این کار نیست و در نتیجه حدود **واقعی** طبقه مد دار، همون (۱۷ و ۱۴) هستن.

توجه کنین که: چون طبقات ما پیوسته هستن، پس فاصله طبقات (I) با عرض طبقات برابر میشه، در نتیجه: $\boxed{I = ۱۱ - ۸ = ۳}$

گام ۳) محاسبه مد:

فاصله (طول) طبقات $\times \dfrac{d_1}{d_1 + d_2} +$ حد پایین **واقعی** طبقه مد دار = **مُد**

$$\boxed{Mo = L_{Mo} + \frac{d_1}{d_1 + d_2} \times I}$$

(ادامه تو فیش بعد)

$$\Rightarrow Mo = ۱۴ + \frac{(۱۸-۱۰)}{(۱۸-۱۰)+(۱۸-۱۲)} \times ۳$$

$$\Rightarrow Mo = ۱۴ + \frac{۸}{۸+۶} \times ۳ \Rightarrow$$

$$۱۴ + \frac{۸}{۱۴} \times ۳ \Rightarrow ۱۴ + \frac{۴}{۷} \times ۳$$

$$\Rightarrow Mo = ۱۴ + \frac{۱۲}{۷} \Rightarrow$$

$$۱۴ + ۱/۷۱ \Rightarrow \boxed{Mo = ۱۵/۷۱ \approx ۱۵/۷}$$

توجه: چون تمام گزینه‌ها تا یک رقم اعشار هستند، پس ما هم باید مقدار ۱۵/۷۱ را تا یک رقم اعشار گرد کنیم که میشه ۱۵/۷.

ایستگاه تغذیه و تناسب اندام:

بعد از شروع به خوردن غذا، حدوداً ۲۰ دقیقه زمان لازم است تا پیام سیری از معده به مغز منتقل شود.

پس آهسته و آرام غذا بخوریم و غذا خوردمان را طول بدهیم.

(مدیریت ۸۲)

در توزیع زیر مُد کدام است؟

C-L	۳-۵	۶-۸	۹-۱۱	**جمع**
F_i	۴	۲۰	۱۲	۳۶

۱) ۶/۳۳

۲) ۶/۵۳

۳) ۷/۵

۴) ۲۰

گام ۱) پیدا کردن طبقه مد دار:

CL	۳ − ۵	۶ − ۸ (طبقه مد دار)	۹ − ۱۱
F_i	۴	$\xleftrightarrow{d_1=۱۶}$ $\boxed{۲۰}$ $\xleftrightarrow{d_2=۸}$	۱۲

گام ۲) پیوسته کردن طبقه مد دار: برای این کار، کافیه که فاصله حد پایین طبقه مد دار (یعنی ۶) تا حد بالای این طبقه (یعنی ۵) رو نصف کنیم $\left(\frac{۶-۵}{۲}=\frac{۱}{۲}=۰/۵\right)$ و این ۰/۵ واحد رو از حد پایین طبقه مد دار کم و به حد بالای طبقه مد دار اضافه کنیم:

$$\begin{cases} \text{۶ :حد پائین} \\ \text{۸ :حد بالا} \end{cases} \Longrightarrow \begin{cases} \text{حد پایین واقعی: } ۶ - ۰/۵ = ۵/۵ \\ \text{حد بالای واقعی: } ۸ + ۰/۵ = ۸/۵ \end{cases}$$

پس:

(۸/۵ و ۵/۵) :حدود واقعی طبقه مد دار

حالا بعد از پیوسته کردن طبقه مد دار، می‌تونیم فاصله طبقات (I) رو هم خیلی راحت بدست بیاریم: $\boxed{I = ۸/۵ - ۵/۵ = ۳}$

(ادامه تو فیش بعد)

گام ۳) محاسبه مد:

$$\text{Mo} = \text{حد پایین واقعی طبقه مد دار} + \frac{d_۱}{d_۱ + d_۲} \times \text{فاصله (طول) طبقات}$$

بنابراین داریم:

$$\Rightarrow Mo = ۵/۵ + \frac{(۲۰-۴)}{(۲۰-۴)+(۲۰-۱۲)} \times ۳$$

$$\Rightarrow Mo = ۵/۵ + \frac{۱۶}{۱۶+۸} \times ۳ \Rightarrow$$

$$۵/۵ + \frac{۱۶ \times ۳}{۲۴} \Rightarrow ۵/۵ + \frac{۴۸}{۲۴}$$

$$\Rightarrow Mo = ۵/۵ + ۲ \Rightarrow \boxed{Mo = ۷/۵}$$

توجه: *گزینه ۴ غلط است، زیرا نشان دهندهٔ فراوانی طبقه مددار است، در حالیکه می‌دانیم، مد یک داده (x_i) است، نه یک فراوانی.*

(ادامه تو پشت فیش)

روش تستی (۵ ثانیه‌ای): مقدار غیردقیق مد برابره با:

$$\boxed{\text{مرکز طبقه مد دار}}: \frac{۶+۸}{۲} = \frac{۱۴}{۲} = ۷$$

CL	طبقه **مجاور**	طبقه مد دار	طبقه **مجاور**
	۳ — ۵	۶ — ۸	۹ — ۱۱
F_i	۴	۲۰	$\boxed{۱۲}$

مقایسه فراوانی‌ها

$۱۲ > ۴$

$\Downarrow$

پس مقدار مد باید از ۷ **به سمت مقادیر بیشتر** (یعنی به سمت طبقه سوم) متمایل و نزدیک بشه $Mo > ۷$ که چنین وضعیتی فقط تو گزینه ۳ دیده میشه: $Mo = ۷ / ۵ > ۷$:**گزینه ۳**

دقت کنین: *اینکه میگیم مقدار مد باید به سمت طبقه سوم نزدیک بشه و مد بزرگتر از ۷ باشه، این بزرگتر بودن باید به حدی باشه که مقدار مد در فاصله حد پایین و بالای طبقه مد دار (۸-۶) قرار داشته باشه. بنابراین یه وقت اشتباهاً نگین گزینه ٤ هم می‌تونه درست باشه، با این استدلال که مقدارش از ۷ بیشتره. چون در گزینه ٤، مقدار مد برابر ۲۰ است و این در حالیه که مقدار مد باید حتماً عددی بین ٦ و ۸ باشه، یعنی بین حد پایین و بالای طبقه مد دار (۸-٦).*

(مهم):

(حسابداری و مدیریت ۸۷)

در جدول توزیع فراوانی دسته‌بندی شده، اگر مد جامعه ۱۶ باشد، فراوانی مطلق دسته‌ی چهارم کدام است؟

حدود دسته	۹-۱۲	۱۲-۱۵	۱۵-۱۸	۱۸-۲۱
فراوانی مطلق	۸	۱۲	۱۵	a

۱) ۷

۲) ۹

۳) ۱۰

۴) ۱۱

جواب: گزینه «۲»

توضیح: چون تو صورت سؤال به ما مقدار مُد رو داده ($Mo = ۱۶$)، پس می‌تونیم از فرمول مد برای بدست آوردن فراوانی مطلق طبقه چهارم استفاده کنیم:

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I}$$

خب، با توجه به جدول زیر می‌بینیم که ما مقدار L_{Mo}، $d_۱$ و $d_۲$ و I رو داریم: $\boxed{d_۲ = ۱۵ - a}$ و $\boxed{d_۱ = ۱۵ - ۱۲ = ۳}$

همچنین چون طبقات ما **پیوسته** است، پس می‌تونیم بگیم که:

اولاً: حد پایین **واقعی** طبقه مد دار، همون $L_{Mo} = ۱۵$ است که در جدول نوشته شده.

ثانیاً: با توجه به **پیوسته** بودن طبقات، فاصله طبقات (I) با عرض طبقات برابره، یعنی: $\boxed{I = ۱۸ - ۱۵ = ۳}$

CL	۹ − ۱۲	۱۲ − ۱۵	۱۵ − ۱۸ (طبقه مد دار)	۱۸ − ۲۱
F_i	۸	۱۲ $\xleftrightarrow{d_۱=۳}$	$\boxed{۱۵}$ $\xleftrightarrow{d_۲=۱۵-a}$	a

(ادامه تو فیش بعد)

توضیح:

چون مقدار مد $(Mo = ۱۶)$ در فاصـــله حد پایین و بالای طبقه **سـوم** (۱۸-۱۵) قرار داره، پس این طبقه (**طبقه سوم**)، طبقه مد داره.

پس:

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I} \Rightarrow$$

$$۱۶ = ۱۵ + \frac{۳}{۳ + (۱۵ - a)} \times ۳ \Rightarrow$$

$$۱۶ = ۱۵ + \frac{۳ \times ۳}{۱۸ - a} \Rightarrow$$

$$۱۶ = ۱۵ + \frac{۹}{۱۸ - a} \Rightarrow$$

$$۱۶ - ۱۵ = \frac{۹}{۱۸ - a} \Rightarrow ۱ = \frac{۹}{۱۸ - a}$$

$$\xrightarrow{\text{طرفین وسطین}} ۱۸ - a = ۹ \Rightarrow ۱۸ - ۹ = a \Rightarrow \boxed{a = ۹}$$

$$\boxed{a = F_۴ = ۹ : \textbf{فراوانی طبقه چهارم}}$$

بدانیم تا بهتر زندگی کنیم:

انجام **تمرینات ورزشی**، سه بار در هفته، استرس و فشار عصبی را کاهش می‌دهد.

با انجام تمرینات ورزشی، مغز ما مواد شیمیایی‌ای ترشح می‌کند (مثل **اندورفین**) که احساس **شادی** در ما ایجاد می‌کند و کمک شایانی به رفع استرس و تنش عصبی ما می‌کند.

در ایامی که برای کنکور درس می‌خونیم، برای شروع تمرینات ورزشی می‌تونیم از روزی ۵ دقیقه شروع کنیم و کم کم اونو به روزی ۱۵ دقیقه برسونیم.

توجه: این کار اصلاً اتلاف وقت نیست، بلکه انرژی‌ای که از این کار در شما ایجاد میشه، باعث میشه مطالعهٔ شما کاراتر و موثرتر انجام بشه.

یعنی ورزش روزانه کوتاه مدت، یه سرمایه گذاری عالیه! پس اونو جدی بگیرین.

(بسیار مهم):

در جدول زیر مقدار مد برابر با ۱۲۳/۵ می‌باشد، مقدار مجهول a کدام است؟

حدود طبقات	۱۰۰-۱۰۹	۱۱۰-۱۱۹	۱۲۰-۱۲۹	۱۳۰-۱۳۹
F_i	۵	۶	a	۴

۱) ۴

۲) ۱۰

۳) ۸

۴) ۱۲

آمار و احتمالات، آموزشگاه مدرسان شریف، ص ۱۷، مثال ۶۵

جواب: گزینه «۳»

اول باید مقدار L_{Mo} ، d_1 و d_2 و I رو بدست بیاریم و اونا رو تو فرمول مد قرار بدیم و از اونجا مقدار a (یعنی فراوانی طبقه سوم) رو بدست بیاریم.

اولاً: چون مقدار مد $(Mo = ۱۲۳/۵)$ در فاصله (۱۲۹-۱۲۰) قرار داره، پس این طبقه (یعنی طبقه سوم)، طبقهٔ مد داره.

ثانیاً: طبقات ما **گسسته** اند، پس ما باید اونا رو پیوسته کنیم (البته برای سرعت عمل بیشتر، **فقط طبقه مد دار** رو پیوسته می‌کنیم):

$$\text{طبقه مد دار}: (۱۲۰-۱۲۹) \xrightarrow{\text{حدود واقعی}} \underset{L_{Mo}}{(۱۱۹/۵-۱۲۹/۵)}$$

L_{Mo} : حد پایین **واقعی** طبقه مد دار

ثالثاً: *بعد از پیوسته شدن طبقات می‌تونیم بگیم که فاصله طبقات (I) با عرض طبقات برابره، پس:* $\boxed{I = ۱۲۹/۵ - ۱۱۹/۵ = ۱۰}$

(ادامه تو فیش بعد)

CL	100-109	110-119	120-129 (طبقه مد دار)	130-139
F_i	۵	۶ $\xleftarrow{d_1=a-6}$	$\boxed{a}$	$\xrightarrow{d_2=a-4}$ ۴

$$\boxed{Mo = L_{Mo} + \frac{d_1}{d_1 + d_2} \times I} \Rightarrow$$

$$۱۲۳/۵ = ۱۱۹/۵ + \frac{a-۶}{(a-۶)+(a-۴)} \times ۱۰ \Rightarrow$$

$$۱۲۳/۵ = ۱۱۹/۵ + \frac{۱۰(a-۶)}{a+a-۶-۴} \Rightarrow$$

$$۱۲۳/۵ = ۱۱۹/۵ + \frac{۱۰a-۶۰}{۲a-۱۰} \Rightarrow$$

$$۱۲۳/۵ - ۱۱۹/۵ = \frac{۱۰a-۶۰}{۲a-۱۰} \Rightarrow$$

$$۴ = \frac{۱۰a-۶۰}{۲a-۱۰} \Rightarrow \xrightarrow{\text{طرفین وسطین}} ۴(۲a-۱۰) = ۱۰a-۶۰ \Rightarrow$$

$$۸a-۴۰ = ۱۰a-۶۰ \Rightarrow -۴۰+۶۰ = ۱۰a-۸a \Rightarrow ۲۰ = ۲a$$

$$a = \frac{۲۰}{۲} = ۱۰ \Rightarrow \boxed{a = F_3 = ۱۰ : \text{فراوانی طبقه سوم (طبقه مد دار)}}$$

(ادامه تو پشت فیش)

نکتهٔ ظریف مصوّر:

تو این ســـؤال، a در واقع نشـــون دهندهٔ فراوانی طبقه مُدداره و ما می‌دونیم که فراوانی طبقهٔ مددار باید از فروانی بقیهٔ طبقات، بیشتر باشه.

پس تو این سـؤال هم، از همـون اول می‌تونستیم بگیم که مقدار a باید از ٦ بزرگتر باشـــه (٦ بزرگترین فراوانی در جدول است)؛

بنابراین با استفاده از همین نکته، براحتی می‌تونستیم بفهمیم که گزینه ۱ $(a = ۴)$ قطعاً غلطه.

در یک جدول فراوانی:

$$L = ۲ \ , \ d_۲ = ۲ \ , \ d_۱ = ۴ \ , \ Mo = ۲ / ۶۷$$

مقدار طول طبقه کدام است؟

۱) ۱

۲) ۲

۳) ۳

۴) ۴

آمار و احتمالات، آموزشگاه مدرسان شریف، ص ۱۱۷، تست ۱۱

پاسخ: گزینه ۲

با توجه به داده‌های مسئله:

$L = ۲$ = حد پایین طبقه مد دار و $Mo = ۲ / ۶۷$: مد

$d_۲ = ۲$ و $d_۱ = ۴$

با توجه به داده‌های بالا، باید از فرمول مد استفاده کنیم:

$$\boxed{Mo = L_{Mo} + \frac{d_۱}{d_۱ + d_۲} \times I} \Rightarrow$$

$$۲ / ۶۷ = ۲ + \frac{۲}{۲ + ۴} \times I \Rightarrow$$

$$۲ / ۶۷ = ۲ + \frac{۲}{۶} \times I \Rightarrow$$

$$۲ / ۶۷ = ۲ + \frac{۱}{۳} \times I \Rightarrow$$

$$۲ / ۶۷ = ۲ + \frac{I}{۳} \Rightarrow$$

$$۲ / ۶۷ - ۲ = \frac{I}{۳} \Rightarrow ۰ / ۶۷ = \frac{I}{۳} \xRightarrow{\text{طرفین وسطین}}$$

طول طبقه: $I = ۳ \times ۰ / ۶۷ \Rightarrow \boxed{I = ۲ / ۰۱ \approx ۲}$

نحوهٔ محاسبه مُد را از طریق **نمودار بافت نگار (هیستوگرام = مستطیلی)** با رسم شکل نشان دهید.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۱۲۶

آمار و کاربرد آن در مدیریت، عادل آذر و منصور مؤمنی، جلد ۱، ص ۱۰۳

«**نمودار بافت نگار (مستطیلی = هیستوگرام)**: ***Histogram chart***»

در رسم نمودار بافت نگار:
*بر روی محور افقی، حدود **واقعی** طبقات (**کرانه**های طبقات) قرار می گیرد*
*و بر روی محور عمودی هم، **فراوانی** (نسبی یا مطلق) قرار می گیرد.*

هر یک از طبقات جدول فراوانی به شکل یک مستطیل نشان داده می‌شوند که:

پهنا (عرض) هر مستطیل نشان دهندهٔ **عرض** اون طبقه است (**یادآوری**: عرض طبقه، یعنی: فاصله حد بالا و پایین طبقه).
و **ارتفاع** هر مستطیل نیز بیانگر **فراوانی** (مطلق یا نسبی) آن طبقه است.

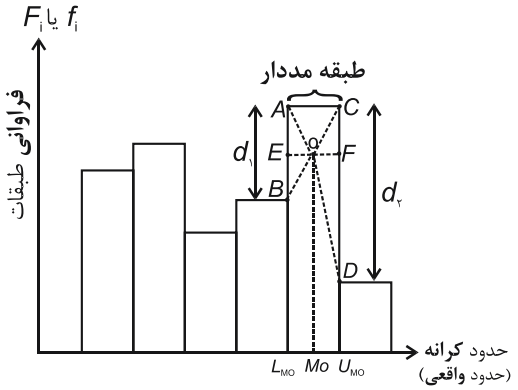
توجه:

از اون‌جایی که طبقات (مستطیل‌ها) را با حدود **واقعی‌شان** نشان می‌دهیم، بنابراین طبقات (مستطیل‌ها) حالت **پیوسته** (بدون فاصله) پیدا می‌کنند، یعنی تمام طبقات (مستطیل‌ها) بهم چسبیده‌اند.

(ادامه تو فیش بعد)

محاسبه مُد از طریق نمودار بافت نگار (هیستوگرام = مستطیلی):

با توجه به اینکه ارتفاع هر مستطیل، همان فراوانی مطلق یا نسبی آن طبقه است، پس **بلندترین** مستطیل، در واقع **بیشترین فراوانی** را داشته و بنابراین **طبقه مد دار** است، یعنی مد داده‌ها بین حد پایین و بالای این طبقه (مستطیل) قرار می‌گیرد:



نکته: چون در رسم بافت نگار از حدود **واقعی** طبقات استفاده می‌کنیم، بنابراین طبقات (مستطیل‌ها) حالت **پیوسته** دارند و در نتیجه همواره:

طول طبقات = **عرض** طبقات = **فاصله** طبقات = I

بنابراین **عرض یا پهنای تمام مستطیل‌ها (تمام طبقات) با هم برابر خواهد بود.**

(ادامه تو پشت فیش)

یافتن مد در شکل فوق (در نمودار هیستوگرام):

بعد از اینکه **طبقه مد دار (بلندترین مستطیل)** را مشخص کردیم، میاییم دوتا خط مورب را از دو گوشهٔ این مستطیل به طبقات مجاورش رسم می‌کنیم. **محل تقاطع این دو خط**، نقطه‌ای را بر روی محور افقی مشخص می‌کند که همان **مد** داده‌ها است.

نحوهٔ ترسیم دو خط مورب:

خط اول را از گوشهٔ سمت راست بلندترین مستطیل (طبقه مد دار) به گوشه سمت راست مستطیل سمت چپ آن می‌کشیم (از نقطه C به B).

خط مورب دوم را نیز از گو شه سمت چپ بلندترین مستطیل به گو شه سمت چپ مستطیل سمت راستی رسم می‌کنیم (از نقطه A به D).

از **محل تقاطع** این دو خط مورب، نقطه‌ای ایجاد می‌شود (نقطه O) که اگر آنرا بر محور افقی عمود کنیم، ***بر روی محور افقی***، مقدار **مُد** مشخص می‌شود.

(**یادآوری**: *مد، یک داده است، نه یک فراوانی؛ پس باید جایش همیشه روی محور افقی باشه، نه روی محور عمودی*).

(ادامه تو فیش بعد)

توجیه هندسی فرمول محاسبه مد از طریق نمودار بافت نگار:

اگر L_{Mo} حد پایین **واقعی** طبقه مد دار و U_{Mo} نیز حد بالای **واقعی** طبقه مد دار باشد و d_1, d_2 نیز به ترتیب اختلاف فراوانی طبقه مد دار را از طبقه **ماقبل** و **مابعد** خود باشد و I هم فاصله طبقه را نشان دهد، می‌خواهیم ثابت کنیم که:

$$\boxed{Mo \approx L_{Mo} + \left(\frac{d_1}{d_1 + d_2}\right) \times I}$$

(**توضیح:** بخاطر سپردن **اثبات** این فرمول الزامی نیست و صرفاً برای آشنایی شما با نحوه بدست آمدن آن، گفته می‌شود):

دو مثلث AOB و COD متشابه هستند، پس طبق قواعد تشابه مثلث‌ها داریم:

$$\frac{\text{قاعده مثلث AOB}}{\text{ارتفاع مثلث AOB}} = \frac{\text{قاعده مثلث COD}}{\text{ارتفاع مثلث COD}} \Rightarrow$$

$$\frac{AB}{OE} = \frac{CD}{OF} \Rightarrow \frac{d_1}{Mo - L_{Mo}} = \frac{d_2}{U_{Mo} - Mo} \Rightarrow$$

با توجه به شکل مشخص است که:

فاصله طبقات + حد پایین طبقه مد دار = حد بالای طبقه مد دار

$$U_{Mo} = L_{Mo} + I$$

(ادامه تو پشت فیش)

بنابراین اگر این مقدار را به جای U_{Mo} در رابطه بـالا قـرار دهـیم، خـواهیم داشت:

$$\xrightarrow{U_{Mo}=L_{Mo}+I} \frac{d_1}{Mo-L_{Mo}} = \frac{d_2}{(L_{Mo}+I)-Mo} \xRightarrow{\text{طرفین وسطین}}$$

$$d_1(L_{Mo}+I-Mo) = d_2(Mo-L_{Mo}) \Rightarrow$$

$$\xrightarrow{d_1 \text{ و } d_2 \text{ را در پرانتزها ضرب می‌کنیم}}$$

$$d_1L_{Mo}+d_1I-d_1Mo = d_2Mo-d_2L_{Mo} \Rightarrow$$

عبارات شامل L_{Mo} *را در سمت چپ و عبارات شامل* Mo *را در سمت راست نگه می‌داریم:* $d_1L_{Mo}+d_2L_{Mo}+d_1I = d_2Mo+d_1Mo$

از Mo *و* L_{Mo} *در طرفین تساوی فوق فاکتور می‌گیریم:*

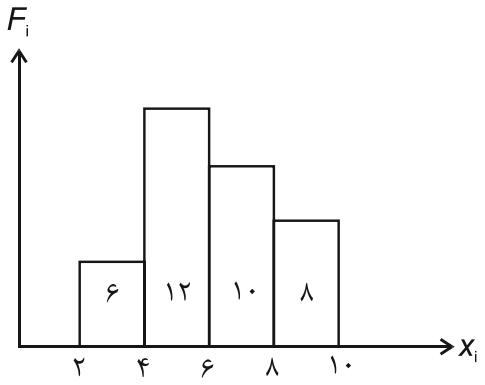
$$L_{Mo}(d_1+d_2)+d_1I = Mo(d_1+d_2) \Rightarrow$$

$$Mo = \frac{L_{Mo}(d_1+d_2)+d_1I}{(d_1+d_2)} \Rightarrow$$

$$Mo = \frac{L_{Mo}\cancel{(d_1+d_2)}}{\cancel{(d_1+d_2)}} + \frac{d_1I}{(d_1+d_2)}$$

$$\Rightarrow \boxed{Mo \approx L_{Mo} + \left(\frac{d_1}{d_1+d_2}\right)\times I}$$

فرض کنید هیستوگرام زیر بدست آمده است. اعداد درون مستطیلها، فراوانی‌های متناظر را نشان می‌دهند. مد کدام است؟



۱) ۵/۵

۲) ۶

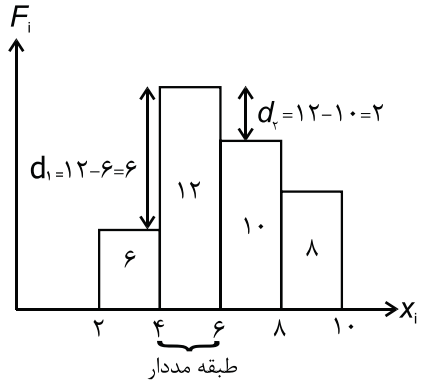
۳) ۱۲

۴) ۴

آمار و احتمالات، آموزشگاه مدرسان شریف، ص ۱۲۰، تست ۴۷

پاسخ: گزینه ۱

۱) راه تستی: با توجه به شکل، طبقه دوم (۶-۴)، طبقه مد دار است چون **پیشترین فراوانی (پیشترین ارتفاع)** رو داره، یعنی مقدار مُد بایستی بین ۴ و ۶ باشد $(۴ < Mo < ۶)$، بنابراین گزینه‌های ۲، ۳ و ۴ حذف می‌شوند و تنها گزینه ۱ صحیح است، زیرا: $۴ < Mo = ۵ / ۵ < ۶$



۲) راه حل محاسباتی:

$$\text{مد دار : Mo} = \text{حد پایین طبقه مد دار} + \frac{d_۱}{d_۱ + d_۲} \times \text{(طول طبقه) فاصله}$$

$$Mo = ۴ + \frac{۶}{۶+۲} \times ۲ \Rightarrow Mo = ۴ + \frac{۶}{\cancel{۲}} \times \cancel{۲} \Rightarrow ۴ + \frac{۶}{۴}$$

$$Mo = ۴ + \frac{۳}{۲} \Rightarrow Mo = ۴ + ۱/۵ \Rightarrow \boxed{Mo = ۵/۵}$$

مزایا و معایب شاخص مرکزی مد:

اصول و روش‌های آماری، فرشادفر، ص ۱۰۳

کاربرد آمار در مدیریت ترافیک، ص ۱۵۰

الف) مزایای نما (مد):

۱- به سادگی قابل محاسبه است (با یافتن داده یا طبقه‌ای که **پیشترین** فراوانی مطلق یا نسبی را دارد).

۲- به آسانی می‌توان آن را از طریق **هندسی (نمودار بافت نگار)** تعیین کرد.

۳- معمولاً از مُد در امور **بازرگانی** استفاده می‌شود.

به عنوان مثال زمانی که یک تاجر لباس مایل است بداند که کدام مارک یا مدل، **پیشترین** فروش را برایش به همراه دارد یا یک فروشنده کفش، مایل است بداند که چه سایزی از کفش‌ها، **پیشترین** فروش را دارند، از شاخص مرکزی **مُد** استفاده می‌کند.

۴- حتی اگر طبقات اول یا آخر، **باز** باشند (در صورتی که مد در این طبقات قرار نداشته باشد)، باز هم **می‌توان** مد را محاسبه کرد.

توجه: اگر مد در یک طبقهٔ **باز** واقع شده باشد، از آنجا که فاصله طبقه باز با طبقه مجاورش نامعلوم است (I نامعلوم است)، پس نمی‌تونیم از فرمول مد برای محاسبه آن استفاده کنیم، در نتیجه مد از طریق فرمول قابل محاسبه **نیست**.

۵- مد را می‌توان برای **همه مقیاس‌ها** اعم از **کمی** یا **کیفی** بدست آورد.

(ادامه تو فیش بعد)

۶- موقعی که داده‌ها به صورت ***فاصله‌ای*** (حد بالا **−** حد پایین) طبقه‌بندی میشن، می‌تونیم بدون نیاز به هیچ محاسبه‌ای بگیم که مقدار مد حتماً ***در حدود طبقه مد دار*** قرار دارد (بین حد پایین و بالای طبقه مد دار).

۷- اگر در مشاهدات، داده‌های فرین (تعداد ***کمی*** مشاهدهٔ ***بسیار کوچک*** یا ***بسیار بزرگ*** نسبت به سایر مشاهدات) وجود داشته باشد، اون وقت مقدار مُد تحت تأثیر بزرگی یا کوچیکی این داده‌ها قرار نمی‌گیره.

به عنوان مثال مد مشاهداتِ ۵ و ۲ و ۲ و ۱ برابر Mo=2 است؛ حال اگر ۲ مشاهدهٔ فرین را به این داده‌ها اضافه کنیم، می‌بینیم که ***مد تحت تأثیر مقدار آنها قرار نمی‌گیرد و تغییر نمی‌کند***:

$$\begin{cases} ۱,۲,۲,۵ \\ Mo = ۲ \end{cases} \longrightarrow \begin{cases} \boxed{-۱۰۰۰},۱,۲,۲,۵,\boxed{۱۰۰۰۰۰} \\ Mo = ۲ \end{cases}$$

نکته:

اما بعداً در بحث شاخص مرکزی میانگین می‌بینیم که شاخص ***میانگین*** (برخلاف مُد) ***به شدت تحت تأثیر مقادیر فرین قرار می‌گیره*** و این یک ***نقطه ضعف*** برای میانگین محسوب می‌شود. (در نتیجه در مواقعی که توزیع داده‌ها، دارای چند مقدار فرین است، شاخص مرکزی مُد بر میانگین ترجیح داده می‌شود.

(ادامه تو پشت فیش)

ب) معایب نما (مد):

۱- مُد بر مبنای کلیه مشـــاهدات نیســـت، یعنی ***در محاســـبه آن از همه داده‌ها استفاده نمی‌شود***، بلکه تنها از اطلاعات مربوط به طبقه مد دار استفاده می‌شود.

۲- بر روی مد، ***نمی‌توان عملیات جبری*** (مانند +، -، × و ÷ و غیره) را انجام داد. و به همین خاطره که **مد کم اهمیت‌ترین و کم ارزش‌ترین شاخص مرکزی است**.

۳- مقدار مد، از یک نمونه به نمونه دیگر، به مقدار قابل توجهی تغییر می‌کند، یعنی ***پراکندگی و تغییرپذیری مد بالا است*** و در نتیجه ***پایداری و ثبات آن کم*** اســـت، به همین علت **مد را** *ناپایدارترین* **شـــاخص مرکزی** قلمداد می‌کنند.

مثلاً اگر جامعه (کل) مشاهدات ما به صورت ۵ و ۴ و ۴ و ۲ و ۲ و ۲ و ۱ باشد و ما چهار نمونه ۴ عضوی از این جامعه انتخاب کنیم، خواهیم داشت:

نمونه اول : ۱,۲,۲,۵ → Mo = ۲

نمونه دوم : ۱,۲,۴,۴ → Mo = ۴

نمونه سوم : ۲,۲,۴,۴ → Mo = ∅

نمونه چهارم : ۱,۲,۴,۵ → Mo = ∅

یعنی مشاهده می‌شود که با تغییر داده‌ها از یک نمونه به نمونه دیگر، مقدار مد دائماً در حال تغییر است، ولی بعداً در ادامه این فصل یاد می‌گیریم که **(ادامه تو فیش بعد)**

مقدار شاخص‌های مرکزی میانه و میانگین، برخلاف مد، از یک نمونه به نمونه‌ای دیگر، به میزان کمتری تغییر می‌کنند، یعنی تغییرپذیری یا پراکندگی آنها نسبت به مد کمتره و برعکس پایداری آنها نسبت به مد بیشتر است:

میانگین < میانه < مد : براساس پراکندگی و تغییرپذیری

(کمترین) (بیشترین)

مد < میانه < میانگین : براساس ثبات و پایداری

(بی ثبات ترین) (پایدار ترین)

۴- برای داده‌های نوع سوم که به صورت فاصله‌ای (C-L) طبقه‌بندی می‌شوند، **نمیشه** مقدار دقیق مُد را بدست آورد، بلکه فقط می‌تونیم مقدار اونو **بطور تقریبی** برآورد کنیم (تخمین بزنیم):

$$Mo \approx L_{\hat{M}o} + (\frac{d_۱}{d_۱ + d_۲})I$$

۵- بعضی وقتا **ممکنه اصلاً مد نداشته باشیم** و بعضی وقتا هم ممکنه همزمان **چند تا مد** داشته باشیم.

مثال: x_i : ۱,۲,۳,۴ $\Rightarrow Mo =$ $\rightarrow$ **مد نداریم**

x_i : ۱,۲,۲,۳,۴,۴,۵,۶,۶ $\Rightarrow \hat{M}o =$ ۲,۴,۶ $\rightarrow$ ۳ تا مد داریم

دانستنی‌های مفید و کاربردی:

درد پاشنه یا کف پا:

برای رهایی از درد پاشنه یا کف پا می‌توانید پای خود را بر روی یك بطری یا میله قرار دهید و آن را روی بطری به جلو و عقب حرکت دهید تا ماساژ داده شود.

با این کار، جریان خون در کف پا بهبود یافته و درد پای شما کاهش می‌یابد.

(اقتصاد ۷۳)

نظر گروهی از سوادآموزان راجع به زمان پخش برنامه نهضت سوادآموزی از سیمای جمهوری اسلامی جمع‌آوری شده است.

کدام شاخص مرکزی برای آن داده‌ها مناسب‌تر است؟

۱) میانگین

۲) میانه

۳) نما

۴) چارک اول

پاسخ: گزینه ۳

نکته: کاربرد مد (نما):

هر گاه در یک جامعه، **بیشترین فراوانی (تکرار)**، معیار سنجش و انتخاب ما باشه (مثلاً هنگام **نظرسنجی**، **رأی‌گیری**، **انتخابات** و ...) اون وقت باید از شاخص مرکزی **مد (نما)** استفاده کنیم.

در این سؤال، با توجه به اینکه معیار ما برای انتخاب بهترین زمان پخش برنامه، براساس **بیشترین** درخواست سوادآموزان است، پس مناسب‌ترین معیار مرکزی برای درخواست‌ها، شاخص **مد (نما)** است.

(بسیار مهم):

«**میانه**: Median»

۱) تعریف شاخص مرکزی **میانه**:

۲) مثال: *در صورتی که میانه نمره دانش‌آموزان یک کلاس ۱۲ باشد، آنگاه نمرات ۵۰ درصد دانش‌آموزان ..... است.*

۱) کمتر یا مساوی ۱۲
۲) مساوی ۱۲
۳) بیشتر از ۱۲
۴) ۱ و ۳

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۵

انواع شاخص‌های مرکزی عبارتند از:

۱- مد	۲- میانه	۳- چندک	۴- میانگین

در فیش‌های قبلی، با شاخص مرکزی مد آشنا شدیم؛ اکنون به بررسی دومین شاخص مرکزی، یعنی میانه می‌پردازیم:

«میانه *(Median)=Md* یا *Me*»

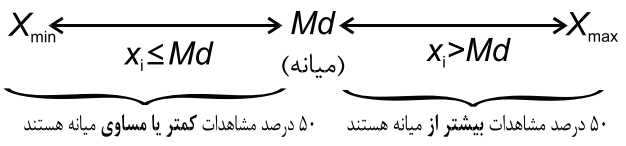
۱) تعریف: اگر داده ها (x های) یک جامعه رو به صورت صعودی (غیرنزولی) مر تب کنیم، اون و قت داده‌ای که **در وسط** قرار می‌گیره و مشاهدات جامعه را به **دو نیمه مساوی تقسیم** می‌کنه، **میانهٔ** مشاهدات خواهد بود. یعنی:

*۵۰ درصد (نصف) مشاهدات، **قبل از میانه** قرار دارن (کوچکتر از میانه‌اند) و*
*۵۰ درصد (نصف) مشاهدات هم **بعد از میانه** قرار دارن (بزرگتر از میانه‌اند).*

توجه: *میانه را به اختصار با نماد* Md یا Me *نشان می‌دهیم.*

نکته بسیار مهم: شاخص میانه (**Md**) جزء ۵۰ در صد **مشاهدات قبل از خودش** است. به عبارت دیگر:

(ادامه تو فیش بعد)



برای درک بهتر نکته فوق به مثال زیر توجه فرمایید:

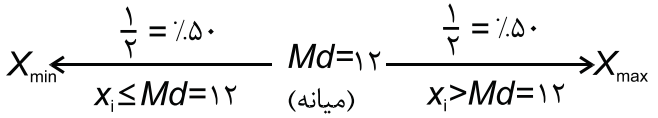
مثال: *در صورتی که میانه نمره دانش‌آموزان یک کلاس ۱۲ باشد، آنگاه نمرات ۵۰ درصد دانش‌آموزان ..... است:*

۱) کمتر یا مساوی ۱۲ ۲) مساوی ۱۲

۳) بیشتر از ۱۲ ۴) ۱ و ۳

پاسخ: گزینه ۴

در صورتی که میانه نمرات دانش‌آموزان یک کلاس Md=۱۲ باشد، به این مفهوم است که اگر نمرات دانش‌آموزان کلاس را به صورت **صعودی (غیرنزولی)** مرتب کنیم، آنگاه:



(ادامه تو پشت فیش)

یعنی ۵۰ درصد نمرات کلاس ***کمتر یا مساوی*** *۱۲ است* (حداکثر برابر ۱۲ است: $x_i \leq ۱۲$) *و ۵۰ درصد نمرات* ***بیشتر از*** *۱۲ است* ($x_i > ۱۲$).

نکته بسیار مهم: *در بیان میانه، فقط می‌توان از عبارت زیر استفاده کرد:*

۱- ۵۰ درصد مشاهدات ***کمتر یا مساوی*** میانه ($\leq$) هستند
(یعنی ۵۰ درصد مشاهدات، **حداکثر** برابر میانه هستند)
۲- ۵۰ درصد مشاهدات ***بیشتر از*** میانه ($<$) هستند.

اما در بیان میانه به هیچ وجه نمی‌توان از عبارات زیر استفاده کرد:

۱- ۵۰ درصد مشاهدات ***بیشتر یا مساوی*** میانه ($\geq$) هستند
(یعنی ۵۰ درصد مشاهدات، **حداقل** برابر میانه هستند).
۲- ۵۰ درصد مشاهدات ***کمتر از*** میانه ($<$) هستند.

علت برقراری روابط فوق در مورد میانه این است که:
میانه جزء ۵۰ درصد مشاهدات قبل از خود (نه بعد از خود) است، بنابراین موقع تفسیر میانه، بجای $>$ باید حتماً از نماد $\leq$ استفاده کنیم.
هم چنین چون میانه جزء ۵۰ درصد مشاهدات بعد از خودش نیست، پس بجای نماد $\geq$ هم باید از نماد $<$ استفاده کنیم (یعنی نباید از علامت مساوی استفاده کنیم).

توجه: *در فیش‌های بعدی، ابتدا با نحوهٔ محاسبه میانه در داده‌های نوع اول (حجم کم) آشنا می‌شویم. سپس به بررسی میانه در داده‌های نوع دوم و سوم می‌پردازیم.*

«نحوهٔ محاسبه میانه در داده‌های نوع اول (با حجم کم)»

مثال: *میانه داده‌های زیر را محاسبه کنید:*

الف) ۶، ۷، ۹، ۰، ۱-

ب) ۹، ۷، ۵ ،۰، ۴ ، ۱-

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۵ و ۲۶

«نحوهٔ محاسبه میانه در داده‌های نوع اول (با حجم کم)»

همان طور که گفته شد، اگر مشاهدات (x_i ها) رو به صورت صعودی (غیرنزولی) مرتب کنیم، آنگاه **دادهٔ وسط**، برابر میانه خواهد بود.

مثال الف) اگه ۵ داده شامل ۶، ۷، ۹، ۰، ۱- داشته باشیم، برای یافتن میانه این داده‌ها، ابتدا باید آنها را **صعودی** مرتب کنیم و سپس **دادهٔ وسطی** را پیدا کنیم:

$$-۱,۰,۹,۷,۶ \xrightarrow{\text{صعودی}} -۱,۰,\boxed{۶},۷,۹$$

دادهٔ وسط = میانه

بنابراین دادهٔ سوم $(x_i = ۶)$ برابر میانه خواهد بود، زیرا از لحاظ مقداری **دقیقاً در وسط** مشاهدات قرار گرفته است، یعنی داده‌های ۰ و ۱- کوچکتر از آن و داده‌های ۷ و ۹ بزرگتر از آن قرار دارند.

مثال ب) اگه داده‌های ما به صورت ۹، ۷، ۵ ، ۰، ۴ ، ۱- باشند، یعنی تعداد داده‌ها **زوج** باشد، آنگاه بعد از مرتب کردن داده‌ها به صورت صعودی، خواهیم دید که بجای یک داده، **دو داده در وسط** مشاهدات قرار می‌گیرند:

$$-۱,۴,۰,۵,۷,۹ \xrightarrow{\text{صعودی}} -۱,۰,۴,۵,۷,۹$$

دو دادهٔ وسط

بنابراین در این حالت (که N زوج است)، باید **متوسط یا میانگین دو دادهٔ وسط** را به عنوان میانه داده‌ها در نظر بگیریم، یعنی:

(ادامه تو فیش بعد)

$$\text{میانه} = Md = \frac{۴+۵}{۲} = \frac{۹}{۲} = ۴/۵ \leftarrow \text{اگر N } \textbf{زوج} \text{ باشد}$$

که در این حالت نیز، سه دادهٔ ۴ و ۰ و ۱- کوچکتر از میانـه (۴/۵) هسـتن و سه دادهٔ ۹، ۷ و ۵ نیز بزرگتر از میانه (۴/۵) هستن، پس در اینجـا هـم مـی-بینیم که میانه در وسط داده‌ها قرار می‌گیرد.

خب، حالا می‌خواهیم توضیحات فوق را به زبان ریاضی نشون بدیم:

۱) اگر n **فرد** باشه:

$$x_۱ \le x_۲ \le \ldots \le \underbrace{x_{\frac{n}{۲}+\frac{۱}{۲}}}_{\textbf{دادهٔ وسط = میانه: Md}} \le \ldots \le x_n$$

۲) اگر n **زوج** باشه:

$$x_۱ \le \ldots \le \underbrace{x_{\frac{n}{۲}} \le x_{\frac{n}{۲}+۱}}_{\textbf{دو دادهٔ وسط=میانه}} \le \ldots \le x_n$$

که در این حالت:

$$\textbf{میانه}: Md = \textbf{میانگین دو دادهٔ وسط} = \frac{x_{\frac{n}{۲}} + x_{\frac{n}{۲}+۱}}{۲}$$

حال برای درک بهتر، می‌خواهیم کاربرد دو رابطه بالا را تو اون دو مثالی که اول این فیش حل کردیم، نشون بدیم:

در مثال (**الف**)، تعداد داده ها **فرد** است $(n = ۵)$ و مشاهدات مـا عبارتنـد از: ۹، ۷، ۶، ۰ و ۱-، پس داریم: **(ادامه تو پشت فیش)**

دادهٔ سوم = محل میانه

$$-۱ < ۰ < \underset{\text{ب}}{۶} < ۷ < ۹$$

دادهٔ وسط = میانه

یعنی در این حالت (که N فرد است)، محل دادهٔ وسط از رابطهٔ زیر بدست می-آید:

$$\textbf{محل دادهٔ وسط} = \frac{n}{۲} + \frac{۱}{۲} \xrightarrow{n=۵} \frac{۵}{۲} + \frac{۱}{۲} = ۳ : \textbf{دادهٔ سوم}$$

یعنی پس از مرتب کردن داده‌ها به صورت صعودی، دادهٔ سوم برابر میانه خواهد بود:

$$\boxed{\textbf{میانه} : Md = x_{\frac{n}{۲}+\frac{۱}{۲}}} \Rightarrow x_{\frac{۵}{۲}+\frac{۱}{۲}} = x_{\frac{۶}{۲}} = \boxed{x_۳ = ۶}$$

اما در مثال **(پ)** که تعداد داده‌ها **زوج** است (N :زوج) و مشاهدات بعد از مرتب شدن به صورت ۹، ۷، ۵، ۴، ، ۰ و ۱- هستن، داریم:

دادهٔ سوم **دادهٔ چهارم**

$$-۱ < ۰ < \underbrace{۴ < ۵}_{\textbf{دو دادهٔ وسط}} < ۷ < ۹$$

یعنی اگر **n زوج** باشد، محل دو دادهٔ وسط از رابطه زیر بدست می‌آید:

$$\boxed{\textbf{محل دو دادهٔ وسط} = \frac{n}{۲}, \frac{n}{۲}+۱} \xrightarrow{n=۶} \frac{۶}{۲}, \frac{۶}{۲}+۱ \Rightarrow ۳, ۴$$

(ادامه تو فیش بعد)

یعنی داده‌های سوم و چهارم در وسط قرار دارند، بنابراین **میانگین** آنها برابر میانه خواهد بود:

$$\boxed{\textbf{میانه} : Md = \frac{x_{\frac{n}{۲}} + x_{\frac{n}{۲}+۱}}{۲}} = \frac{x_۳ + x_۴}{۲} = \frac{۴+۵}{۲} = \frac{۹}{۲} = ۴/۵$$

با توجه به توضیحات فوق، می‌توان مراحل محاسبه میانه در داده‌های نوع اول را به صورت زیر بیان نمود:

گام ۱) مرتب کردن داده‌ها بصورت **صعودی (غیرنزولی)** و کدگذاری آنها از ۱ تا N.

گام ۲) یافتن **محل میانه:**

داده **وسط** = محل میانه ⟵ اگه N **فرد** باشه

دقیقاً **بین ۲ دادهٔ وسط** = محل میانه ⟵ اگه N **زوج** باشه

گام ۳) مشخص کردن مقدار میانه از بین داده‌های صعودی مرتب شده

ایستگاه ورزش:

بهترین شیوه برای گرم کردن بدن پیش از شروع تمرینات ورزشی، پیاده روی آرام به مدت ۵ دقیقه است.

توجه:

انجام حرکات کششی پیش از فعالیت اصلی ورزشی و پیش از گرم شدن کامل بدن، ممکن است به ماهیچه‌ها و تاندون‌های ما آسیب برساند.

یادآوری:

انجام ۱۰-۱۵ دقیقه ورزش روزانه،دارای فواید زیر است:

۱.کمک به **بهبود یادگیری** (بواسطه اکسیژن رسانی بیشتر به مغز)

۲. **افزایش انرژی و حوصله** شما برای مطالعه (بواسطه ترشح هورمونهای نشاط بخش از جمله اندورفین)

۳.تغییر روحیه شما و **رهایی شما از استرس و نگرانی**

(اقتصاد ۸۳)

برای مقادیر متغیر تصادفی **X** *به شرح زیر:*

x : ۶۰,۴۰,۶۰,۸۰

۱) فقط میانگین حسابی و میانه با هم برابرند.

۲) فقط میانگین حسابی و نما با هم برابرند.

۳) میانگین حسابی، میانه و نما با هم برابرند.

۴)) فقط میانه و نما با هم برابرند.

پاسخ: گزینه ۳

گام ۱) صعودی مرتب کردن داده‌ها و کدگذاری آنها از ۱ تا N:

کد $\Rightarrow$ (۱) (۲) (۳) (۴)

x_i : ۴۰ , $\underbrace{۶۰ \ , \ ۶۰}_{\text{دو دادهٔ وسط}}$, ۸۰

گام ۲) تعیین محل میانه:

چون تعداد داده‌ها **زوج** است (N=۴)، **دو** مشاهده (دوم و سوم) در وسط قرار می‌گیرند و بنابراین میانه، برابر با **میانگین دو دادهٔ وسط** خواهد بود:

$$Md = \frac{x_۲ + x_۳}{۲} = \frac{۶۰+۶۰}{۲} = \frac{۱۲۰}{۲} \Rightarrow \boxed{Md = ۶۰}$$

خب، حالا میریم سراغ میانگین:

$$\text{میانگین حسابی} = \mu = \frac{\text{جمع مشاهدات}}{\text{تعداد مشاهدات}} = \frac{\sum x_i}{N} = \frac{۴۰+۶۰+۶۰+۸۰}{۴}$$

$$\Rightarrow \mu = \frac{۲۴۰}{۴} \Rightarrow \boxed{\mu = ۶۰}$$

و حالا نوبت مُده: $\boxed{Mo = ۶۰}$: مُد

همانطور که می‌بینیم مقدار هر سه پارامتر مرکزی میانگین حسابی، میانه و نما (مد) برابر ۶۰ شده است: $\boxed{\mu = Md = Mo = ۶۰}$

(ادامه تو فیش بعد)

راه حل تستی:

تو این ســـؤال داده‌های ما حالت **متقارن** دارن، یعنی دو دادهٔ ۶۰ و ۶۰ در وســـط توزیع قرار دارن و دو دادهٔ دیگه (۴۰ و ۸۰) هر کدوم **با فاصـــله مساوی** از دادهٔ وسط، در دو طرفش قرار گرفته‌اند:

۲۰ واحد فاصله **۲۰ واحد فاصله**

x_i : ۴۰ , ۶۰ , ۶۰ , ۸۰

دو دادهٔ وسط

به این جور توزیع‌ها، توزیع **متقارن** می‌گیم.

نکته تســـتی: بعدها یاد می‌گیریم که ***در توزیع‌های متقارن، همه شاخص‌های مرکزی با هم برابرن،*** یعنی:

$\mu = Md = Mo$: در توزیع **متقارن**

پس تو این سؤال هم:

داده وسط $= ۶۰ = Mo = Md = \mu \Rightarrow$ توزیع **متقارن**

ایستگاه معنویت:

وضو:
کسی که وضو می‌گیرد و با حوله خشک کند، یک حسنه برایش نوشته می‌شود و کسی که وضو بگیرد و صبر کند تا دست و رویش خودبخود خشک شود، سی حسنه برایش نوشته می‌شود.
کسی که برای غیر نماز، وضوی خود را تجدید کند (دوباره وضو بگیرد)، خداوند ، توبه او را بدون استغفار تجدید می‌کند.
امام صادق (علیه السلام)

«نحوهٔ محاسبه میانه در داده‌های نوع دوم (حجم زیاد، تنوع کم) با داشتن فراوانی مطلق»

مثال : میانه در جدول زیر کدام است؟

x_i	۴	۱	۰	۷
F_i	۳	۸	۵	۹

۱) ۷/۵ ۲) ۶/۵

۳) ۴ ۴) ۱

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۷ و ۲۸

محاسبه میانه در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با در اختیار داشتن فراوانی مطلق

در این حالت برای محاسبه میانه به صورت زیر عمل می‌کنیم:

الف) مرتب کردن x_i ها (طبقات جدول) به صورت **صعودی**.
ب) یافتن **محل** میانه: $C_{Md} = \frac{N}{۲} + \frac{۱}{۲}$: **محل میانه**
ج) محاسبه فراوانی **تجمعی** طبقات F_{C_i}: $F_{C_i} = F_i + F_{C_{i-۱}}$
د) یافتن **طبقه میانه دار** (طبقه‌ای که میانه در آن قرار دارد): *اولین طبقه جدول از چپ به راست ، که در آن* $F_{C_i} \geq \frac{N}{۲} + \frac{۱}{۲}$ *باشد.*

تو این فیش و فیشای بعد، می‌خوایم مراحل بالا را در قالب ۳ مثال متفاوت بیان کنیم:

در مثال ۱ تعداد داده‌ها **فرده** $n = ۲۵$

در مثال ۲ تعداد داده‌ها **زوجه** $n = ۲۶$

در هر دو مثال فوق، میانه **در یک طبقه** (طبقه میانه دار) قرار می‌گیرد، ولی در مثال ۳، میانه **در فاصلهٔ بین دو طبقه** قرار می‌گیرد.

بدین ترتیب شما با انواع حالات ممکن در محاسبه میانه آشنا خواهید شد:

(ادامه تو فیش بعد)

حل مثال: **پاسخ: گزینه ۴**

x_i	۴	۱	۰	۷	
F_i	۳	۸	۵	۹	$N = ۲۵$

الف) صعودی مرتب کردن طبقات (x_i ها):

(این کارو تو جدول پشت این فیش انجام داده‌ایم).

ب) یافتن **محل** میانه:

محل میانه : $C_{Md} = \frac{N}{۲} + \frac{۱}{۲} \xrightarrow{N = \sum F_i = ۲۵}$

۳(اُمین داده: $C_{Md} = \frac{۲۵}{۲} + \frac{۱}{۲} = ۱۲/۵ + ۰/۵ = ۱۳$: **محل میانه**

یعنی دادهٔ ۱۳اُمی، میانه است.

ج) محاسبه فراوانی **تجمعی** طبقات (F_{C_i}):

که در جدول پشت فیش مشخص شده.

د) یافتن **طبقه میانه‌دار: اولین** طبقه‌ای که در آن $F_{C_i} \geq \frac{N}{۲} + \frac{۱}{۲}$ است: $(F_{C_i} \geq ۱۳)$، یعنی طبقه **دوم**: $(x = ۱)$ طبقه میانه داره:

(ادامه تو پشت فیش)

(دادهٔ ۶اُم تا ۳۱اُم) (دادهٔ اول تا ۵اُم)

x_i **صعودی**	۰	۱	۴	۷	
F_i	۵	۸	۳	۹	$\sum F_i = N = ۲۵$
F_{C_i}	۵	(۸+۵) $\boxed{۱۳}$	(۳+۱۳) ۱۶	(۹+۱۶) ۲۵	

همان طور که در جدول بالا مشخص شده است:

از داده اول تا ۵اُم در طبقه اول $(x = ۰)$ *قرار گرفته*

و از دادهٔ ۶اُم تا ۳۱اُم در طبقه دوم $(x = ۱)$

کد: ۱ ۲ ۳ ۴ ۵ ۶ ۷ ۸ ۹ ۱۰ ۱۱ ۱۲ ۱۳

$$x_i = \underbrace{۰ , ۰ , ۰ , ۰ , ۰}_{\text{طبقه } \textbf{اول}} , \underbrace{۱ , ۱ , ۱ , ۱ , ۱ , ۱ , ۱ , \boxed{۱}}_{\text{طبقه } \textbf{دوم}} , ۴ , ۴ , ۴ , ...$$

بنابراین میانه (دادهٔ ۱۳اُم) در طبقه دوم قرار می‌گیرد، پس طبقه دوم $(x = ۱)$ **طبقه میانه دار** *است. بنابراین میانه برابر* $\boxed{x = ۱}$ *است، زیرا دقیقاً در وسط مشاهدات قرار می‌گیرد (۱۲ مشاهده قبلش و ۱۲ مشاهده نیز بعدش قرار میگیرن).*

«نحوهٔ محاسبه میانه در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با داشتن فراوانی مطلق»

توضیح: در مثال ۱ (در فیش قبل)، تعداد داده‌ها **فرد** بود $(n = ۲۵)$، اما حالا در مثال ۲، حالتی را برر سی می‌کنیم که تعداد داده‌ها **زوج** ا ست $(n = ۲۶)$:

مثال ۲: *میانه در جدول زیر کدام است؟*

x_i	۴	۱	۰	۷
F_i	۸	۴	۵	۹

۱) ۶/۵ ۲) ۴

۳) ۰/۵ ۴) ۱

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۷ و ۲۸

پاسخ: گزینه ۲

x_i	۴	۱	۰	۷	
F_i	۸	۴	۵	۹	$N=۲۶$

گام ۱) مرتب کردن طبقات (x_i ها) به صورت صعودی (غیرنزولی)

(به جدول فیش بعد توجه کنین)

گام ۲) یافتن محل میانه:

$$C_{Md}=\frac{N}{۲}+\frac{۱}{۲}\xrightarrow{N=\sum F_i=۲۶}$$: **محل** میانه

$$C_{Md}=\frac{۲۶}{۲}+\frac{۱}{۲}=۱۳+۰/۵=۱۳/۵$$ **۱۳/۵(اُمین داده:**

یعنی میانه، دادهٔ ۱۳/۵اُمی است. یعنی چون تعداد داده‌ها **زوج** است، پس **دو** دادهٔ ۱۳اُم و ۱۴اُم در وســـط قرار می‌گیرند، بنابراین میانه در **بین این دو داده** قرار می‌گیرد (یعنی ۱۳/۵اُمین داده).

گام ۳) یافتن طبقه میانه دار: از چپ به راست، اولین طبقه‌ای که در آن فراوانی تجمعی **بزرگتر یا مساوی** $C_{Md}=۱۳/۵$ **است** $(F_{C_i}\geq\frac{N}{۲}+\frac{۱}{۲}=۱۳/۵)$**، یعنی طبقه سوم** $(x=۴)$ **را به عنوان طبقه میانه دار انتخاب می‌کنیم.** **(ادامه تو فیش بعد)**

کد: ۱ ۲ ۳ ۴ ۵ ۶ ۷ ۸ ۹ ۱۰ ۱۱ ۱۲ ۱۳ ۱۴ ۱۵ ۱۶ ۱۷

x_i: ۰ , ۰ , ۰ , ۰ , ۰ , ۱ , ۱ , ۱ , ۱ , ۴ , ۴ , ۴ , $\boxed{۴}$, $\boxed{۴}$, ۴ , ۴ , ۴ , ۷,۷,...

طبقه **اول** طبقه **دوم** طبقه **سوم**

همون طور که می‌بینین، چون داده ۱۳/۵ اُمی (میانه) **بین** دو دادهٔ ۱۳ اُم $(x = ۴)$ و ۱٤ اُم $(x = ۴)$ است، پس میانه **در وسط آنها** یعنی **بین** $x_{۱۴} = ۴ , x_{۱۳} = ۴$ قرار می‌گیره، یعنی در واقع میانه، میانگین این دو دادهٔ وسط است:

گام ۴) محاسبه میانه:

$$\text{Md}: \textbf{میانه} = \textbf{میانگین دو دادهٔ وسط} = \frac{x_{۱۳} + x_{۱۴}}{۲} = \frac{۴+۴}{۲} = \frac{۸}{۲} = ۴$$

روش سریعتر:

تو گام ۲ فهمیدیم که میانه، ۱۳/۵ اُمین داده است، یعنی میانه دقیقاً در وسط دادهٔ ۱۳ اُم و ۱٤ اُم قرار داره.

با نگاه به جدول می‌بینیم که از دادهٔ ۱۰ اُم تا ۱۷ اُم، همگی در طبقه سوم $(x = ۴)$ قرار دارن، پس نتیجه می‌گیریم که داده‌های ۱۳ اُم و ۱٤ اُم هم در همین طبقه $(x = ۴)$ قرار دارن و بنابراین مقدار هر دوی اونا مساوی ٤ است. در نتیجه نقطه وسط اونا هم حتماً مساوی ٤ خواهد شد، یعنی میانه مساوی ٤ است ،به همین راحتی، به همین خوش مزگی!!!

(ادامه تو پشت فیش)

(دادهٔ اول تا ۵ اُم) (دادهٔ ۱۰ اُم تا ۱۷ اُم)

(۶ اُم تا ۹ اُم)

x_i صعودی	۰	۱	$\boxed{۴}$	۷	
F_i	۵	۴	۸	۹	$\sum F_i = N = ۲۶$
F_{C_i}	۵	(۴+۵) $\boxed{۹}$	(۸+۹) ۱۷	(۹+۱۷) ۲۶	

یادآوری:

برای محاسبه فراوانی **تجمعی** هر طبقه (F_{C_i})، فراوانی **مطلق** آن طبقه (F_i) را با فراوانی **تجمعی طبقه قبلش** $(F_{C_{i-۱}})$ جمع می‌کنیم:

$$F_{C_i} = F_i + F_{C_{i-۱}} \xrightarrow{\text{مثلاً در طبقه ۲}} F_{C_۲} = F_۲ + F_{C_۱} = ۴ + ۵ = ۹$$

توجه: در این مثال و مثال فیش قبل (مثال ۱)، ما حالتی را بررسی کردیم که میانه دقیقاً **در یک طبقه** (x_i) قرار می‌گرفت (**نه بین دو طبقه**) ولی تو فیش بعد (در مثال ۳) حالتی را بررسی می‌کنیم که میانه **بین** دو طبقه (**بین دو** x_i **متمایز**) قرار می‌گیره.

«نحوهٔ محاسبه میانه در داده‌های نوع دوم (حجم زیاد، تنوع کم) با داشتن فراوانی مطلق»

مثال ۳: *میانه در جدول زیر کدام است؟*

x_i	۴	۱	۰	۷
F_i	۳	۸	۵	۱۰

۱) ۴ ۲) ۰/۵

۳) ۲/۵ ۴) ۶

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۷ و ۲۸

پاسخ: گزینه ۳

x_i	۴	۱	۰	۷	
F_i	۳	۸	۵	۱۰	$N=۲۶$

گام ۱) صعودی کردن طبقات (x_i ها): جدول صعودی شده در فیش بعد رسم شده.

گام ۲) یافتن محل میانه:

$$\textbf{محل میانه}: C_{Md}=\frac{N}{۲}+\frac{۱}{۲} \xrightarrow{N=\sum F_i=۲۶}$$

$$C_{Md}=\frac{۲۶}{۲}+\frac{۱}{۲}=۱۳/۵ \Rightarrow$$ میانه ۱۳/۵اُمین داده است.

کد: ۱ ۲ ۳ ۴ ۵ ۶ ۷ ۸ ۹ ۱۰ ۱۱ ۱۲ ۱۳ ۱۴ ۱۵ ۱۶

x_i: ۰ , ۰ , ۰ , ۰ , ۰ , ۱ , ۱ , ۱ , ۱ , ۱ , ۱ , ۱ , $\boxed{۱}$, $\boxed{۴}$, ۴ , ۴ , ۷,۷,...

طبقه اول — طبقه دوم — طبقه سوم

گام ۳) یافتن طبقه میانه دار:

میانه، ۱۳/۵اُمین دادهٔ ماست، یعنی میانه **دقیقاً در وسط** دادهٔ ۱۳اُم و ۱۴اُم قرار داره. خب، با نگاه به جدول فیش بعد می‌بینیم که:

(ادامه تو فیش بعد)

اولاً: دادهٔ ۱۳ اُم در طبقه دوم (x=۱) قرار داره، پس مقدار دادهٔ ۱۳ اُم مساوی ۱ است: $\boxed{x_{۱۳} = ۱}$

ثانیاً: دادهٔ ۱٤ اُم در طبقه سوم ($x = ۴$) قرار داره، پس مقدارش مساوی ٤ است: $\boxed{x_{۱۴} = ۴}$

خب، پس اینکه گفتیم میانه دقیقاً در وسط داده ۱۳اُم و ۱٤اُم قرار داره، یعنی میانه دقیقاً در وســط $x = ۱$, $x = ۴$ قرار داره، یعنی برای محا سبه میانه، باید میانگین $x = ۱$, $x = ۴$ (یعنی همون نقطه وسط $x = ۱$, $x = ۴$) رو بدست بیاریم:

(اول تا ۵اُم) (۶اُم تا ۱۳اُم) (۱۴اُم تا ۱۶اُم)

x_i صعودی	۰	۱	۴	۷	
F_i	۵	۸	۳	۱۰	$\sum F_i = N$
F_{C_i}	۵	(۸+۵) ۱۳	(۳+۱۳) ۱۶	(۱۰+۱۶) ۲۶	

گام ۴) محاسبه مقدار میانه:

$$\textbf{میانه}: Md = \frac{x_{۱۳} + x_{۱۴}}{۲} = \frac{۱ + ۴}{۲} = \frac{۵}{۲} = ۲/۵ \Rightarrow \boxed{Md = ۲/۵}$$

نجوا با خداوند:

خدایا تا وقتی شما پشتم هستید، دیگر مهم نیست چه کسانی جلویم ایستاده‌اند،

تا وقتی تکیه ام به شماست، احساس کوهی را دارم که زلزله هم تکانش نمی‌دهد،

تا وقتی دســـت یاری شـــما را حس می‌کنم، معنی بی کســـی و بی‌پناهی را نمی‌فهمم،

و تا وقتی به من اجازه می‌دهید دوست‌تان باشم، دیگر برایم مهم نیست چه کسانی برای رفاقت دست رد به سینه‌ام می‌زنند.

(حسابداری ۸۲)

جمعیت خانواده‌های یک روستا به صورت زیر است:

جمعیت خانواده	۱	۲	۳	۴	۵	۶	جمع
تعداد	۵	۱۰	۴۰	۲۵	۱۵	۵	۱۰۰

میانه جمعیت خانواده، چقدر است؟

۱) ۳

۲) ۳۲/۵

۳) ۳/۵

۴) ۴

جواب: گزینه «۱»

گام ۱) صعودی کردن طبقات (x_i ها): تو این سؤال، به طور خودبخود، طبقات ما حالت صعودی دارن.

(۱ تا ۵ اُمی) (۶ تا ۵۵ اُمی) (۱۶ اُم تا ۵۵ اُمی)

x_i صعودی	۱	۲	۳	۴	۵	۶	
F_i	۵	۱۰	۴۰	۲۵	۱۵	۵	$\sum F_i = N = ۱۰۰$
F_{C_i}	۵	(۱۰+۵) ۱۵	(۴۰+۱۵) $\boxed{۵۵}$				

گام ۲) یافتن محل میانه:

$$C_{Md} = \frac{N}{۲} + \frac{۱}{۲} \Rightarrow \frac{۱۰۰}{۲} + \frac{۱}{۲} = ۵۰ + ۰/۵ = ۵۰/۵ \ : \textbf{محل میانه}$$

یعنی میانه، ۵۰/۵ اُمین داده اســـت و این بدین مفهومه که میانه دقیقاً در وسط دو دادهٔ ۵۰اُم و ۵۱اُم است.

خب، حالا با نگاه به جدول (و سطر F_{C_i}) می‌بینیم که دادهٔ ۱۶ تا ۵۵اُم در طبقه ســـوم $(x = ۳)$ قرار دارن، پس مســـلماً دادهٔ ۵۰/۵ اُم هم باید در همین طبقه $(x = ۳)$ قرار داشـــته باشـــه، بنابراین مقدار دادهٔ ۵۰/۵اُم هم حتماً مساوی ۳ است، یعنی:

میانه: دادهٔ ۵۰/۵اُم $\Longrightarrow$ طبقه سوم $\Rightarrow \boxed{Md = x = ۳}$

(ادامه تو فیش بعد)

توجه: روش دیگه برای بدست آوردن میانه، اینه که بگیم:

اولین طبقه‌ای که فراوانی **تجمعی**اش **بزرگتر یا مساوی** $\frac{N}{۲}+\frac{۱}{۲}=۵۰/۵$ باشه: $(F_{C_i} \geq ۵۰/۵)$، میانه ماست.

خب، با نگاه به جدول می‌بینیم که: اولین طبقه‌ای از جدول که در اون $F_{C_i} \geq ۵۰/۵$ است، طبقه سومه $(x=۳)$، پس میانه برابره با:

$$F_{C_۳}=۵۵ \geq ۵۰/۵ \xrightarrow{\text{طبقه سوم}} \boxed{Md=x=۳}$$

دانستنیهای جالب و کاربردی:

۱) برای درخشندگی مو یک قاشق سرکه به موهای خود زده، سپس آب بکشید.

۲) برای پاک کردن آدامس از روی لباس لباس را به مدت یک ساعت در فریزر قرار دهید.

۳) برای جلوگیری از ریزش اشک هنگام پوست کندن پیاز آدامس بجوید.

۴) برای جوشاندن سریع تخم مرغ به آب آن نمک اضافه کنید.

۵) برای امتحان تازگی تخم مرغ آن را در آب بگذارید اگر به صورت افقی قرار گرفت تازه است.
اگر به صورت کج قرار گرفت ۳-٤ روزه است.
و اگر عمودی قرار گرفت ۱۰ روزه است.
و اگر برروی سطح آب ایستاد، کهنه است.

(حسابداری ۸۳)

میانه داده‌های جدول زیر کدام است؟

x_i	۵	۱۲	۱۵	۲۰
f_i	۶	۸	۱۲	۴

۱) ۱۰

۲) ۱۳/۵

۳) ۱۵

۴) هیچکدام

جواب: گزینه «۳»

توجه: در این ســـؤال منظور از f_i در واقع همان فراوانی **مطلق** (F_i) است، چون:

اولاً: این فراوانی‌ها اعداد صحیح و بزرگتر از یک هستند.

ثانیاً این فراوانی‌ها به صورت اعشاری (یعنی بین صفر و یک) نیستن، پس نمی‌تونن فراوانی نسبی باشن.

گام ۱) صعودی کردن طبقات: طبقات ما خودشون حالت صعودی دارن.

(۷اُم تا ۴(اُم) (۲۴ اُم تا ۲۶ اُم)

x_i صعودی	۵	۱۲	۱۵	۲۰	
F_i	۶	۸	۱۲	۴	$\sum F_i = N = ۳۰$
F_{C_i}	۶	۸+۶ ۱۴	۱۲+۱۴ $\boxed{۲۶}$		

گام ۲) یافتن محل میانه:

$$\boxed{\textbf{محل میانه}: C_{Md} = \frac{N}{۲} + \frac{۱}{۲}} \Rightarrow \frac{۳۰}{۲} + \frac{۱}{۲} = ۱۵ + ۰/۵ = ۱۵/۵$$

یعنی میانه، داده **۱۵/۵**(اُمی است. **(ادامه تو فیش بعد)**

خب، با نگاه به جدول بالا می‌بینیم که از دادهٔ ۱۵اُم تا ۲۶اُم در طبقه سوم $(x = ۱۵)$ قرار دارن، پس نتیجه می‌گیریم که دادهٔ ۱۵/۵اُمی هم تو همین طبقه $(x = ۱۵)$ قرار داره، یعنی:

میانه: دادهٔ ۱۵/۵اُم $\xrightarrow{\text{طبقه سوم}} \Rightarrow \boxed{Md = x = ۱۵}$

راه حل دوم:

میانه برابره با **اولین** طبقه‌ای که فراوانی **تجمعیش** بزرگتر یا مساوی $\frac{N}{۲} + \frac{۱}{۲}$ است:

$$\boxed{F_{C_i} \geq \frac{N}{۲} + \frac{۱}{۲}} \rightarrow F_{C_i} \geq ۱۵/۵ \xrightarrow{\text{طبقه سوم}} F_{C_۳} = ۲۶ \geq ۱۵/۵$$

یعنی اولین طبقه‌ای که F_{C_i} اون بزرگتر یا مساوی ۱۵/۵ است، طبقه سومه $(x = ۱۵)$، پس: $\boxed{Md = x = ۱۵}$

امام علی علیه السلام
لا ادَبَ مَعَ غَضَبٍ
غررالحکم - ح ۱۰۵۲۹
با خشم ، تربیت ممکن نیست.
چون خشم و غضب حماقتی مُزمن نیست
حسد و قهر و غضب، از صفت مُومن نیست
فرمود امام عارفان، حضرت عشق
در حالت خشم، تربیت مُمکن نیست
4Ohadis.vcp.ir

«نحوهٔ محاسبه میانه در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با در اختیار داشتن فراوانی نسبی»

مثال: *میانه در جدول داده‌های زیر کدام است؟*

x_i	۴	۱	۰	۷
f_i	۰/۳	۰/۱	۰/۴	۰/۲

۱) ۴

۲) ۲/۵

۳) ۱

۴) صفر

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۲۸

«نحوهٔ محاسبه میانه در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با داشتن فراوانی نسبی»

در این حالت برای محاسبه میانه به صورت زیر عمل می‌کنیم:

گام ۱) مرتب کردن طبقات (x_iها) به صورت ***صعودی (غیرنزولی)***

گام ۲) محاسبه فراوانی ***تجمعی نسبی*** طبقات (f_{C_i}):

$$f_{C_i} = f_i + f_{C_{i-1}}$$

گام ۳) *یافتن طبقه میانه دار: اولین* طبقه‌ای که فراوانی تجمعی نسبی آن ***بیشتر از*** $\frac{1}{2} = 0/5$ باشد ($f_{C_i} \geq 0/5$)، بیانگر میانه داده‌ها است، زیرا میانه مشاهده‌ای است که ۵۰ درصد یا ۰/۵ مشاهدات قبل از آن قرار دارند.

مثال: میانه در جدول داده‌های زیر کدام است؟

x_i	۴	۱	۰	۷
f_i	۰/۳	۰/۱	۰/۴	۰/۲

۱) ۴ ۲) ۲/۵

۳) ۱ ۴) صفر

پاسخ: گزینه ۳ *(ادامه تو فیش بعد)*

توجه: *چون فراوانی‌ها بصورت اعشاری‌اند و جمع آنها برابر یک است، پس بیانگر فراوانی نسبی هستند.*

گام ۱) ابتدا طبقات (x_i ها) را **به ترتیب صعودی** مرتب می‌کنیم (به صورت جدول زیر):

گام ۲) فراوانی **تجمعی نسبی** طبقات را بدست می‌آوریم:

x_i **صعودی**	۰	$\boxed{۱}$	۴	۷	
f_i	۰/۴	۰/۱	۰/۳	۰/۲	$\sum f_i = ۱$
f_{C_i}	۰/۴	$\boxed{۰/۵}$	۰/۸	۱	

گام ۳) اولین طبقه‌ای که در آن $f_{C_i} \geq ۰/۵$ باشد، طبقه دوم ($x_۲ = ۱$) است. بنابراین میانه در این طبقه قرار دارد و مقدار آن برابر $x = ۱$ است: $\boxed{Md = ۱}$

ایستگاه تغذیه:

برای داشتن تغذیه سالم باید باهوش باشیم و با آگاهی انتخاب کنیم.

*مثلاً برای دریافت کلسیم الزامی ندارد خودمان را مجبور به مصرف شیر کنیم؛ اگر شیر را دوست نداریم، می توانیم همین کلسیم را با مصرف **بادام، کلم** یا حتی **استخوان ماهی** جبران کنیم.*

(مهم):

«نحوهٔ محاسبه میانه در داده‌های نوع سوم (حجم زیاد، تنوع زیاد)

با داشتن فراوانی مطلق»

مثال ۱ (طبقات پیوسته): *در جدول زیر میانه برابر است با:*

حدود	۱-۵	۵-۹	۹-۱۳	۱۳-۱۷
F_i	۵	۱۰	۳	۲

۱) ۵ ۲) ۷

۳) ۹ ۴) ۱۳

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۲۹

نحوهٔ محاسبه میانه در داده‌های نوع سوم (حجم زیاد، تنوع زیاد) با داشتن فراوانی مطلق

در مورد داده‌های نوع ســـوم، به دلیل تنوع زیاد آنها، امکان در نظر گرفتن یک طبقه مجزا برای هر نوع داده (x_i) وجود نداره، بنابراین در این حالت میایم داده‌ها را در قالب طبقاتی که دارای حد پایین و بالا (L-U) هستند، به صورت **فاصله‌ای یا پیوسته** طبقه‌بندی می‌کنیم.

مراحل محاسبه میانه برای این نوع از داده‌ها به صورت زیره:

گام ۱) محاسبه فراوانی **تجمعی** طبقات: $F_{C_i} = F_i + F_{C_{i-1}}$

گام ۲) یافتن **محل میانه (طبقه میانه دار):** اولین طبقه‌ای از چپ به راست که فراوانی **تجمعی**اش بیشتر یا مساوی $\frac{N}{2}$ باشد $(F_{C_i} \geq \frac{N}{2})$ طبقه میانه دار است.

گام ۳) یافتن **مقدار** میانه از رابطه زیر:

$$\text{میانه} = \text{حد پایین واقعی طبقه میانه دار} + \left(\frac{\frac{N}{2} - \text{فراوانی تجمعی طبقه قبل}}{\text{فراوانی مطلق طبقه میانه دار}}\right) \times \text{طول طبقه (فاصله طبقه)}$$

$$\Rightarrow Md = L_{Md} + \frac{\frac{N}{2} - F_{C_{i-1}}}{F_{Md}} \times I$$

(ادامه تو فیش بعد)

L_{Md}: حد پایین **واقعی** طبقه میانه دار

$F_{C_{i-۱}}$: فراوانی **تجمعی** طبقه قبل از طبقه میانه دار

F_{Md}: فراوانی **مطلق** طبقه میانه دار

I: فاصله طبقات (طول طبقات) است که در طبقات پیوسته، با عرض طبقات نیز برابر است.

تذکر: اگر طبقات **گسسته** باشند، بایستی بعد از اجرای گام ۲، یعنی بعد از یافتن طبقه میانه دار، آن را **پیوسته** کنیم. یعنی باید حدود **واقعی** طبقه میانه دار را بدست بیاریم و بعدش گام (۳) را انجام دهیم.

برای درک بهتر مراحل فوق ۲ مثال را ذکر می‌کنیم که مثال ۱، مربوط به طبقات پیوسته و مثال ۲ مربوط به طبقات گسسته است که مربوط به فیش بعدی است:

مثال ۱ (طبقات پیوسته): **پاسخ: گزینه ۲**

گام ۱) محاسبه فراوانی تجمعی طبقات: $\boxed{F_{C_i} = F_i + F_{C_{i-۱}}}$

گام ۲) یافتن طبقه میانه دار: اولین طبقه‌ای که در آن $F_{C_i} \geq \frac{N}{۲} = \frac{۲۰}{۲} = ۱۰$ باشه، طبقه (۹-۵) است، پس این طبقه، طبقه میانه داره.

((ادامه تو پشت فیش))

توجه: چون طبقات ما **پیوسته** هستن، پس دیگه نیازی به یافتن حدود واقعی طبقه میانه دار نیست.

گام ۳) یافتن مقدار میانه از رابطه زیر:

Md= حد پایین **واقعی** طبقه میانه دار + $\frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i}$ × فاصله (طول) طبقات

$$Md = L_{Md} + \frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i} \times I$$

$$Md = ۵ + \frac{\frac{۲۰}{۲} - ۵}{۱۰} \times ۴ \Rightarrow$$

$$Md = ۵ + \frac{۱۰ - ۵}{۱۰} \times ۴ \Rightarrow ۵ + \frac{\cancel{۵}}{\cancel{۱۰}} \times ۴ \Rightarrow$$

$$Md = ۵ + \frac{۱}{۲} \times ۴ \Rightarrow ۵ + ۲ \Rightarrow \boxed{Md = ۷}$$

حدود	۱-۵	۵-۹	۹-۱۳	۱۳-۱۷	
F_i	۵	۱۰	۳	۲	$\sum F_i = N = ۲۰$
F_{C_i}	۵	۱۵	۱۸	۲۰	

(مهم):

«نحوهٔ محاسبه میانه در داده‌های نوع سوم (حجم زیاد، تنوع زیاد)

با داشتن فراوانی مطلق»

مثال ۲ (طبقات گسسته): *در جدول زیر، مقدار میانه کدام است؟*

حدود طبقات	۵-۸	۹-۱۲	۱۳-۱۶	۱۷-۲۰
F_i	۵	۱۲	۱۴	۱۰

۱) ۱۴ ۲) ۱۳

۳) ۱۳/۵ ۴) ۱۴/۲

آمار و احتمالات، آموزشگاه مدرسان شریف،ص ۱۵

پاسخ: گزینه ۳

گام ۱) محاسبه فراوانی تجمعی طبقات (به صورت جدول زیر).

گام ۲) یافتن محل میانه (طبقه میانه دار): اولین طبقه‌ای که در آن $F_{C_i} \geq \frac{N}{۲} = \frac{۴۱}{۲} = ۲۰/۵$ برقرار است، طبقه (۱۶-۱۳) است، و این طبقه، طبقه میانه داره.

توجه: با توجه به گسسته بودن طبقات، قبل از انجام گام سوم، باید اول طبقه میانه دار را پیوسته کنیم:

گام ۳) یافتن مقدار میانه:

$$\boxed{Md = L_{Md} + \frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i} \times I} \Rightarrow$$

$$Md = ۱۲/۵ + \frac{\frac{۴۱}{۲} - ۱۷}{۱۴} \times ۴ \Rightarrow ۱۲/۵ + \frac{۲۰/۵ - ۱۷}{۱۴} \times ۴ \Rightarrow$$

$$Md = ۱۲/۵ + \frac{۳/۵}{۱۴} \times ۴ \Rightarrow ۱۲/۵ + \frac{۱۴}{۱۴} \Rightarrow$$

$$\boxed{\text{میانه} : Md = ۱۲/۵ + ۱ = ۱۳/۵}$$

CL	۵-۸	۹-۱۲	۱۳-۱۶	۱۷-۲۰	
F_i	۵	۱۲	۱۴	۱۰	$\sum F_i = N = ۴۱$
F_{C_i}	۵	۱۷	۳۱	۴۱	

(مدیریت ۷۳)

میانه داده‌های جدول زیر کدام است؟

CL (فاصله طبقات)	۲۰-۲۹	۳۰-۳۹	۴۰-۴۹
F_i (فراوانی)	۳	۶	۷

۱) ۳۴/۶

۲) ۳۴/۵

۳) ۳۷/۸

۴) ۳۷/۳

پاسخ: گزینه ۳

گام ۱) محاسبه فراوانی تجمعی طبقات (به صورت جدول زیر):

CL	۲۰-۲۹	طبقه میانه دار ۳۰ – ۳۹	۴۰-۴۹	
F_i	۳	$\boxed{۶}$	۷	$\sum F_i = N = ۱۶$
F_{C_i}	$\boxed{۳}$	۹	۱۶	

گام ۲) یافتن محل میانه (طبقه میانه دار):

$$F_{C_i} \geq \frac{N}{۲} = \frac{۱۶}{۲} = ۸ \Rightarrow$$

$\xrightarrow{\text{طبقه دوم}} F_{C_۲} = ۹ \geq ۸ \Rightarrow$ پس طبقه دوم، طبقه میانه دار است

گام ۳) پیوسته کردن طبقه میانه دار: کافیه ۰/۵ واحد از حد پایین کم و ۰/۵ واحد به حد بالای طبقه میانه دار اضافه کنیم:

$$(۳۰ - ۳۹) \xrightarrow{\pm ۰/۵} \boxed{(۲۹/۵ - ۳۹/۵)}$$

گام ۴) محاسبه مقدار میانه:

$$\boxed{Md = L_{Md} + \frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i} \times I}$$

$$Md = ۲۹/۵ + \frac{\frac{۱۶}{۲} - ۳}{۶} \times ۱۰ \Rightarrow ۲۹/۵ + \frac{۸ - ۳}{۶} \times ۱۰ \Rightarrow$$

$$Md = ۲۹/۵ + \frac{۵}{۶} \times ۱۰ \Rightarrow ۲۹/۵ + \frac{۵۰}{۶} \Rightarrow ۲۹/۵ + ۸/۳ \Rightarrow \boxed{Md = ۳۷/۸}$$

بسیار مهم:

(حسابداری و مدیریت ۸۶)

میانه در توزیع آماری ۵۰ مشاهده دسته‌بندی شده برابر ۴۱ می‌باشد. اگر طول دسته‌ها ۵، فراوانی طبقه میانه دار ۱۰ و مجموع فراوانی‌های ماقبل طبقه میانه دار برابر ۱۸ باشد، حدود دسته میانه دار کدام است؟

۱) (۴۱/۵ و ۳۶/۵)

۲) (۴۲ و ۳۷)

۳) (۴۲/۵ و ۳۷/۵)

۴) (۴۳ و ۳۸)

پاسخ: گزینه ۳

گام ۱) نوشتن داده‌های مسئله:

تعداد داده‌ها : $N = ۵۰$

میانه : $Md = ۴۱$

طول دسته : $I = ۵$

فراوانی طبقه میانه دار : $F_i = F_{Md} = ۱۰$

فراوانی تجمعی دسته ماقبل میانه : $F_{C_{i-۱}} = F_{C_{Md-۱}} = ۱۸$

حدود دسته میانه دار : $(L_{Md} - U_{Md}) = ?$

گام ۲) استفاده از فرمول میانه:

$$\boxed{Md = L_{Md} + \frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i} \times I} \Rightarrow$$

$$۴۱ = L_{Md} + \frac{\frac{۵۰}{۲} - ۱۸}{۱۰} \times ۵ \Rightarrow$$

$$۴۱ = L_{Md} + \frac{۲۵ - ۱۸}{۱۰} \times ۵ \Rightarrow \quad ۴۱ = L_{Md} + \frac{۷ \times ۵}{۱۰} \Rightarrow$$

$$۴۱ = L_{Md} + \frac{۳۵}{۱۰} \Rightarrow ۴۱ = L_{Md} + ۳/۵ \Rightarrow ۴۱ - ۳/۵ = L_{Md} \Rightarrow$$

(ادامه تو فیش بعد)

حد پایین **واقعی** طبقه میانه دار: $L_{Md} = ۳۷ / ۵$

خب، چون فاصله حد پایین هر طبقه از حد بالای خودش به اندازه طول طبقات (I) است، پس برای بدست آوردن حد **بالای** طبقه میانه دار، کافیه که حد پایین این طبقه رو به طول طبقات (I=۵) اضافه کنیم:

حد بالای **واقعی** طبقه میانه دار: $U_{Md} = L_{Md} + I = ۳۷ / ۵ + ۵ = ۴۲ / ۵$

بنابراین:

حدود **واقعی** طبقه میانه دار: $(۳۷ / ۵, ۴۲ / ۵)$

ایستگاه سلامت:

کیسه های نایلونی و نون داغ!

شاید شماها هم از اون دسته افرادی باشید که موقعی که بعد از خریدن نان، اونو تو کیسه نایلونی قرار میدن!

مشکلی که اینجاست اینه که نون داغ با ترکیبات داخل نایلون، واکنش های شیمیایی مضری میره که باعث سرطان یا مشکلات کبدی میشه.

خب، حالا راه حل چیه؟

بهترین راه اینه که نون رو که خریدیم اول بذاریم کمی خنک بشه و بعد اونو داخل پارچه سفید تمیز و بدون چاپ و نوشته حمل کنیم (بهتره از پارچه رنگی یا روزنامه هم استفاده نکنیم، چون رنگهای اونا به نون جذب میشه و دوباره همون داستان قبلی پیش میاد).

با خودمون مهربون تر باشیم.

بسیار مهم:

(برنامه ریزی شهری ۸۷)

در ۶۰ مشاهده دسته‌بندی شده، میانه ۲۳، فاصله طبقات ۳ و فراوانی طبقه میانه دار ۱۲ و فراوانی تجمعی طبقه میانه دار ۳۸ می‌باشد. حدود دسته میانه دار کدام است؟

۱) (۲۴ و ۲۱)

۲) (۲۵ و ۲۲)

۳) (۲۴/۵ و ۲۱/۵)

۴) (۲۵/۵ و ۲۲/۵)

پاسخ: گزینه ۲

گام ۱) نوشتن داده‌های مسئله:

تعداد داده‌ها: $N = ۶۰$

میانه: $Md = ۲۳$

طول دسته: $I = ۳$

فراوانی طبقه میانه دار: $F_i = F_{Md} = ۱۲$

فراوانی **تجمعی** طبقه میانه دار: $F_{C_i} = F_{C_{Md}} = ۳۸$

حدود دسته میانه دار: $(L_{Md} - U_{Md}) = ?$

گام ۲) استفاده از فرمول میانه:

میانه: $$\boxed{Md = L_{Md} + \frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i} \times I}$$

برای استفاده از فرمول فوق، ابتدا باید $F_{C_{i-۱}}$ (فراوانی تجمعی طبقه ماقبل میانه دار) را پیدا کنیم:

= فراوانی تجمعی یک طبقه

فراوانی تجمعی طبقه ماقبلش + فراوانی مطلق اون طبقه

$$\Rightarrow F_{C_i} = F_i + F_{C_{i-۱}}$$

(ادامه تو فیش بعد)

پس می‌تونیم بنویسیم:

$$۳۸ = ۱۲ + F_{C_{i-۱}} \Rightarrow F_{C_{i-۱}} = ۳۸ - ۱۲ \Rightarrow \boxed{F_{C_{i-۱}} = ۱۶}$$

خب، حالا تازه می‌تونیم بریم سراغ فرمول میانه:

$$۲۳ = L_{Md} + \frac{\frac{۶۰}{۲} - ۲۶}{۱۲} \times ۳ \Rightarrow$$

$$۲۳ = L_{Md} + \frac{۳۰ - ۲۶}{۱۲} \times ۳ \Rightarrow ۲۳ = L_{Md} + \frac{۴}{۱۲} \times ۳$$

$$\Rightarrow ۲۳ = L_{Md} + \frac{۴ \times ۳}{۱۲} \Rightarrow ۲۳ = L_{Md} + \frac{۱۲}{۱۲} \Rightarrow$$

$$۲۳ = L_{Md} + ۱ \Rightarrow L_{Md} = ۲۳ - ۱ \Rightarrow$$

حد پایین **واقعی** طبقه میانه دار: $\boxed{L_{Md} = ۲۲}$

گام ۴) نوشتن حدود طبقه میانه دار:

اگه به حد پایین طبقه میانه دار، فاصـــله طبقات $(I = ۳)$ رو اضافه کنیم به حد بالای طبقه میانه دار می‌رسیم:

$$۲۲ + ۳ = ۲۵ \Rightarrow$$ حد **بالای** واقعی طبقه میانه دار: $U_{Md} = L_{Md} + I$

پس: (۲۲,۲۵) : حدود **واقعی** طبقه میانه دار

ایستگاه تغذیه:

«مصرف کنسروها را محدودتر کنیم»

کنسروها حاوی مواد نگهدارنده زیادی هستند.

(**دقت کنیم که:** این نگهدارنده‌ها در واقع یک نوع سم ضعیف شده هستند تا مانع رشد میکروبها و قارچها شوند، ولی باید بدانیم که مصرف این نگهدارنده‌ها ***برای ما هم ضرر دارد*** و آسیب شدیدی به کبد ما می‌رساند).

پس بیاییم تا جایی که می توانیم از مصرف کنسروها بپرهیزیم یا اینکه فقط در شرایط خاص و موقتی از آنها استفاده کنیم.

تا جایی که امکان دارد از مواد غذایی نگهداری شده در شیشه استفاده کنیم تا کنسروهای فلزی. به خاطر این که امکان دارد عناصر شیمیایی موجود در فلزات به محتویات کنسروها نیز نفوذ کنند.

توجه: بعد از باز کردن کن سرو آن را بلافا صله م صرف کنید. مدت زمان نگهداری کنسرو باز شده نباید بیشتر از ۴۸ ساعت باشد، در غیر اینصورت موجب مسمومیت شدید و حتی مرگ می‌شود.

«محاسبه میانه در داده‌های نوع سوم (حجم زیاد، تنوع زیاد)

با داشتن فراوانی نسبی»

مثال: *با توجه به جدول طبقه‌بندی شده زیر، مقدار میانه کدام است؟*

CL	۱۰-۲۰	۲۰-۳۰	۳۰-۴۰	۴۰-۵۰
f_i	۰/۱	۰/۴	۰/۳	۰/۲

۱) ۲۵

۲) ۳۰

۳) ۲۰

۴) ۴۰

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۰

«محاسبه میانه در داده‌های نوع سوم (حجم زیاد، تنوع زیاد)

با داشتن فراوانی نسبی»

مراحل محاسبه میانه بشرح زیر است:

گام۱) محاسبه **فراوانی تجمعی نسبی** طبقات: $f_{c_i} = f_i + f_{c_{i-1}}$

گام۲) یافتن محل میانه (طبقه میانه‌دار): اولین طبقه‌ای که در آن $f_{C_i} \geq \frac{1}{2} = 0/5$ برقرار است، طبقه میانه‌دار خواهد بود.

تذکر مهم: *در صورت گسسته بودن طبقات، قبل از انجام گام۳، اول باید طبقه میانه‌دار را پیوسته کنیم، یعنی باید حدود **واقعی** طبقه میانه‌دار را بدست آوریم.*

گام۳) یافتن **مقدار میانه** از رابطه زیر:

$$\text{میانه} = \text{حد پایین واقعی طبقه میانه‌دار} + \left(\frac{\left(\frac{1}{2} - \text{فراوانی تجمعی نسبی طبقه قبل میانه‌دار}\right)}{\text{فراوانی نسبی طبقه میانه‌دار}}\right) \times \text{طول طبقه}$$

$$\Rightarrow Md = L_{Md} + \frac{\frac{1}{2} - f_{C_{i-1}}}{f_i} \times I$$

نتیجه مهم: اگر در جدول، بجای فراوانی **مطلق** (F_i)، فراوانی **نسبی** (f_i) را داشتیم، برای محاسبه میانه به همان روش فراوانی مطلق عمل کرده و تنها:

(ادامه تو فیش بعد)

بجای $\frac{N}{۲}$ در فرمول، از $\frac{۱}{۲} = ۰/۵$ استفاده می‌کنیم.

و بجای $F_{C_{i-۱}}$، از $f_{C_{i-۱}}$ و بجای F_i هم از f_i استفاده می‌کنیم.

حل مثال: پاسخ: گزینه ۲

C-L	۱۰-۲۰	۲۰-۳۰	۳۰-۴۰	۴۰-۵۰
f_i	۰/۱	۰/۴	۰/۳	۰/۲

تذکر: قبل از حل سؤال، ابتدا بایستی **نوع فراوانی** را در جدول تشخیص دهیم:

با توجه به اینکه فراوانی‌ها اعشاری‌اند و جمع آنها برابر با یک است: $(\sum f_i = ۱)$

پس با فراوانی‌های **نسبی** سر و کار داریم.

گام۱) محاسبه فراوانی **تجمعی نسبی** (در جدول پشت این فیش).

گام۲) یافتن محل میانه (طبقه میانه‌دار): اولین طبقه‌ای که در آن، $f_{C_i} \geq \frac{۱}{۲} = ۰/۵$ است، طبقه (۳۰-۲۰) است، پس این طبقه، طبقه میانه داره و در نتیجه مقدار میانه در محدودهٔ حد بالا و پایین این طبقه (بین ۲۰ و ۳۰) قرار داره.

گام۳) یافتن مقدار میانه:

$$Md = L_{Md} + \frac{۰/۵ - f_{C_{i-۱}}}{f_i} \times I$$

(ادامه تو پشت فیش)

$$Md = ۲۰ + \frac{۰/۵ - ۰/۱}{۰/۴} \times ۱۰ \Rightarrow ۲۰ + \frac{۰/۴}{۰/۴} \times ۱۰ \Rightarrow$$

$$Md = ۲۰ + (۱ \times ۱۰) \Rightarrow ۲۰ + ۱۰ \Rightarrow \boxed{Md = ۳۰}$$

نکته: قبلاً در مورد مد گفتیم که مقدار مد باید حتماً بین حد پایین و بالای طبقه مد دار قرار داشته باشد.
بطور مشابه در مورد میانه هم می‌تونیم بگیم که مقدار میانه هم باید حتماً بین حد پایین و بالای طبقه میانه‌دار باشد.
در مثال فوق نیز $Md = ۳۰$ در محدوده (۳۰-۲۰) قرار داره.

توجه کنین که: با توجه به نکته بالا، می‌تونستیم از همون اول، گزینه ۴ $(Md = ۴۰)$ رو حذف کنیم، چون مقدارش بین حد پایین و بالای طبقه میانه‌دار (۳۰-۲۰) قرار نداره.

C-L	۱۰-۲۰	طبقه میانه دار: $\boxed{۲۰} - ۳۰$	۳۰-۴۰	۴۰-۵۰	
f_i	۰/۱	$\boxed{۰/۴}$	۰/۳	۰/۲	$\sum f_i = ۱$
f_{C_i}	$\boxed{۰/۱}$	۰/۵	۰/۸	۱	

تذکر مهم: **در اینجا نیز چنانچه طبقات گسسته باشند، پس از یافتن محل میانه (طبقه میانه‌دار)، ابتدا باید طبقه میانه‌دار را پیوسته کنیم و بعدش بریم سراغ گام۳، یعنی محاسبه میانه.** **(ادامه تو فیش بعد)**

راه حل تستی باحال:

نکته تستی: *اگر در سطر فراوانی نسبی تجمعی، مستقیماً خود عدد ۰/۵ رو ببینیم، اون وقت تنها کافیه که حد* **بالای** *اون طبقه رو به عنوان* **میانه** *در نظر بگیریم.*

مثلا در این سؤال، ما مستقیماً می‌تونیم فراوانی نسبی تجمعی $f_{C_i} = ۰/۵$ رو در ستون مربوطه به طبقه دوم (۳۰-۲۰)، ببنیم، پس در این حالت بدون نیاز به هیچ محاسبه‌ای می‌تونیم بگیم که:

میانه = حد بالای طبقه مورد نظر $\Rightarrow Md = ۳۰$

دانستنیهای کاربردی و مفید:

برای خلاص شدن از دست حشرات در هنگام شب برگهای نعنا را نزدیک تخت و بالش و در اطراف اطاق خود قرار دهید.

(مهم):

«**مشخصات میانه** (Md)»

در مورد مشـــخصـــات شـــاخص مرکزی میانه به ســـؤالات زیر پاسخ دهید:

۱. یک جامعه آماری چندتا میانه، چندتا مد و چندتا میانگین می‌تونه داشته باشه؟

۲. در صورتی که به بزرگترین عدد یک سری داده، عدد ثابتی را اضافه کنیم یا از کوچکترین عدد یک سری داده، عدد ثابتی را کم کنیم، میانه چه تغییری خواهد نمود؟ چرا؟

۳. برخلاف مد (نما)، که تابع .....(۱)..... بود، میانه تابع .....(۲)..... است، به عبارت دیگر تا زمانی که تغییرات در داده‌ها، .....(۳)..... را تغییر ندهد، مقدار میانه تغییر نمی‌کند.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۱
آمار و احتمالات، آموزشگاه مدرسان شریف، ص ۱۶
آمار کاربردی ۱، عالِم تبریز و احمد هژبر، ص ۴۹

مشخصات میانه $(Me = Md)$:

۱. هر جامعه آماری **فقط یک میانه** دارد، یعنی میانه **منحصر بفرد** است و **همیشه** وجود دارد، زیرا در هر نوع جامعه‌ای با تعداد عضوهای فرد یا زوج، **همیشه** می‌تونیم **یک نقطه** را در وسط داده‌ها (به عنوان میانه آن داده‌ها) مشخص کنیم.

به عبارت دیگر هیچ وقت داده‌ها، دارای ۲ یا چندین میانه نخواهند بود و هیچ وقت هم پیش نمی‌آید که داده‌ها، میانه نداشته باشند.

اما در مورد مد (نما)، همان طور که قبلاً یاد گرفتیم، جامعهٔ ما ممکن است **۲ یا چندین مد** داشته باشند، یعنی مد برخلاف میانه منحصر بفرد نیست و همچنین مد برخلاف میانه ممکن است اصلاً وجود نداشته باشد (مثل زمانی که فراوانی همه مشاهدات با هم برابر است و در نتیجه مدی وجود نخواهد داشت).

میانگین از این لحاظ شبیه میانه است، یعنی هر جامعه آماری **فقط یک میانگین (حسابی)** دارد، یعنی میانگین **منحصر بفرد** است.

علاوه بر آن، میانگین مانند میانه و برخلاف مد، **همیشه** وجود دارد، زیرا میانگین (حسابی) در واقع همان معدل (متوسط) داده‌ها است، یعنی همیشه می‌توان داده‌ها را با هم جمع زد و بر تعدادشان تقسیم کرد و میانگین حسابی شون رو بدست آورد: **(ادامه تو فیش بعد)**

$$\text{میانگین} : \frac{\sum X_i}{N}$$

۲. اگر به بزرگترین عدد یک سری داده، عدد ثابتی را **اضافه** کنیم، از کوچکترین عدد یک سری داده، عدد ثابتی را **کم** کنیم، چون با این کار **ترتیب داده‌ها تغییر نمی‌کند**، پس این تغییرات، هیچ تاثیری روی میانه نداشته و در نتیجه مقدار آن تغییری نمی‌کند.

مثال:

$$\begin{cases} \boxed{۵},۸,۱۲,۱۷,\boxed{۳۴} \\ \text{میانه} = Md = ۱۲ \end{cases} \xrightarrow[x_{max}=۳۴ \xrightarrow{۱۷} ۴۱]{x_{min}=۵ \xrightarrow{۳} ۲} \begin{cases} \boxed{۲},۸,۱۲,۱۷,\boxed{۴۱} \\ Md = ۱۲ \end{cases}$$

همان‌طور که دیده می‌شود کم کردن عدد ۳ از کمترین مشاهده $(X_{min} = ۵)$ و اضافه کردن عدد ۷ به بزرگترین مشاهده $(X_{max} = ۳۴)$ ***تغییری در ترتیب داده‌ها ایجاد نکرده*** و در نتیجه مقدار میانه $(Md = ۱۲)$ نیز تغییر نمی‌کند.

(ادامه تو پشت فیش)

۳. با توجه به ویژگی ۲ که در بالا به آن اشـــاره کردیم، می‌تونیم بگیم که:

> برخلاف مد (نما) که تابع **فراوانی (بیشترین فراوانی)** است، میانه تابع **ترتیب داده‌ها** است.

به عبارت دیگر تا زمانی که تغییرات در داده‌ها، ترتیب داده‌ها را تغییر ندهد، مقدار میانه تغییر نمی‌کند.

مثلاً در مشاهدات زیر، میانه برابر ۷ است؛ حال اگر همه داده‌ها بجز میانه تغییر کند، ولی ترتیب مشاهدات تغییر نکند (یعنی همچنان عدد ۷ در وسط مشاهدات قرار گیرد)، آنگاه میانه این مشاهدات جدید برابر با مشاهدات قبلی خواهد بود، یعنی میانه ***تغییری نمی‌کند:***

{۳,۴,۵,[۷],۹,۶۰,۷۰ ; **میانه** $= Md = ۷$} $\xrightarrow[\text{بجز داده وسط}]{\text{تغییر داده‌ها}}$ {۱,۲,۲,[۷],۷,۷,۸ ; **میانه** $Md = ۷$}

اما مد، تابع **فراوانی** داده‌ها است، مثلاً :

{۱,۲,۲,۲,۷,۷,۷,۸ ; **مد** $= Mo = ۲,۷$} $\xrightarrow{\text{تغییر فراوانی داده‌ها}}$ {۱,۲,۲,۷,۷,۷,۸ ; **مد** $= Mo = ۷$}

(مهم):

«مشخصات میانه (ادامه)»

۴. هرگاه کل داده‌ها را در عدد ثابتی ضرب یا تقسیم کنیم، میانه چه تغییری می‌کند؟

۵. آیا اگر کل داده‌ها را با عدد ثابتی جمع یا از آن کم کنیم، میانه داده‌ها تغییر می‌کند؟ با مثال توضیح دهید.

۶. آیا اگر کل داده‌ها را در عدد ثابتی ضرب یا تقسیم کنیم و با عدد ثابتی جمع یا از آن کم کنیم، میانه تغییر خواهد کرد؟ با مثال توضیح دهید.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۱
آمار و احتمالات، آموزشگاه مدرسان شریف، ص ۱۶
آمار کاربردی ۱، عالِم تبریز و احمد هژبر،ص ۴۹

۴. هرگاه کل داده‌ها (نه بخشی از آنها) را در عدد ثابتی (مانند a) ضرب یا بر آن عدد تقسیم کنیم، **میانه نیز به همان نسبت تغییر خواهد کرد :**

$$\begin{cases} x_i : x_1, x_2, \ldots, x_n \\ \text{میانه } Md_x \end{cases} \xrightarrow{y_i = ax_i} \begin{cases} y_i : ax_1, ax_2, \ldots, ax_n \\ \boxed{\text{میانه : } Md_y = aMd_x} \end{cases}$$

یعنی اگر کل داده‌ها را در عدد ثابت a ضرب شوند، میانه دادها نیز در عدد ثابت a ضرب می‌شود.

زیرا در این حالت، ترتیب داده‌ها تغییر نمی‌کند و تنها داده و سط (میانه) و نیز بقیه داده‌ها، a برابر می‌شـــوند، مثلاً اگر تمامی داده‌ها در عدد ثابت ۲- ضرب شوند، میانه آنها هم در ۲- ضرب می‌شود:

۱,۲,$\boxed{۳}$,۴,۵ $\xrightarrow{\text{تمام داده‌ها} \times -۲}$ ۲ -, ۴ -, $\boxed{۶ -}$, ۸ -, ۱۰ -

میانه = ۳ $\xrightarrow{\text{(میانه هم در ۲- ضرب میشه)}}$ میانه = ٦-

(ادامه تو فیش بعد)

۵. اگر کل داده‌ها را با عدد ثابتی (مانند b) جمع و یا از آن عدد کم کنیم، میانه داده‌ها نیز به همان نسبت تغییر خواهد کرد:

$$\begin{cases} x_i : x_1, x_2, \dots, x_n \\ \textbf{میانه}: Md_x \end{cases} \xrightarrow{y_i = x_i \pm b} \begin{cases} y_i : x_1 \pm a, \dots, x_n \pm a \\ \textbf{میانه}: \boxed{Md_y = Md_x \pm a} \end{cases}$$

مثال: $x_i : ۱,۲,\boxed{۳},۴,۵,$ $\xrightarrow{y_i : x_i - ۳}$ $y_i : -۲,-۱,\boxed{۰},۴,۵$

$\downarrow$ ۳= میانه $\xrightarrow{\textbf{(میانه هم منهای ۳ میشه)}}$ $\downarrow$ ۰= میانه

---

۶. با توجه به ویژگی‌های ۴ و ۵ در بالا، اگر هر دوی این ویژگی‌ها را با هم ترکیب کنیم، یعنی:

اگر کل داده‌ها را در عد ثابتی (مانند α) ضرب یا تقسیم کنیم و نیز با عدد ثابتی (مانند b) جمع یا از آن کم کنیم، میانه داده‌ها نیز به همان نسبت تغییر خواهد کرد:

$$\begin{cases} x_i : x_1, x_2, \dots, x_n \\ \textbf{میانه}: Md_x \end{cases} \xrightarrow{y_i = ax_i \pm b} \begin{cases} y_i : ax_1 \pm b, \dots, ax_n \pm b \\ \textbf{میانه}: \boxed{Md_y = aMd_x \pm b} \end{cases}$$

(ادامه تو پشت فیش)

مثال: *اگر میانه مشاهدات* $x_1, x_2, \ldots, x_6$ *برابر ۴ باشد، آنگاه میانه مشاهدات* $-2x_1+3, -2x_2+3, \ldots, -2x_6+3$ *برابر است با:*

۱) ۸-　　۲) ۴

۳) ۵-　　۴) ۷

پاسخ: گزینه ۳

چون تمام داده‌ها (x_i ها) ابتدا در ۲- ضرب شده و سپس با ۳ جمع شده‌اند، میانه داده‌ها نیز ابتدا در ۲- ضرب و سپس با ۳ جمع می‌شود:

$$\begin{cases} x_i = x_1, x_2, \ldots, x_6 \xrightarrow{y_i = -2x_i + 3} \\ \text{میانه}: Md = 4 \xrightarrow{Md_y = -2Md_x + 3} \end{cases} \begin{cases} y_i : -2x_1+3, \ldots, -2x_6+3 \\ \text{میانه}: Md_y = -2(4)+3 = -5 \end{cases}$$

(مهم):

«مشخصات میانه (ادامه)»

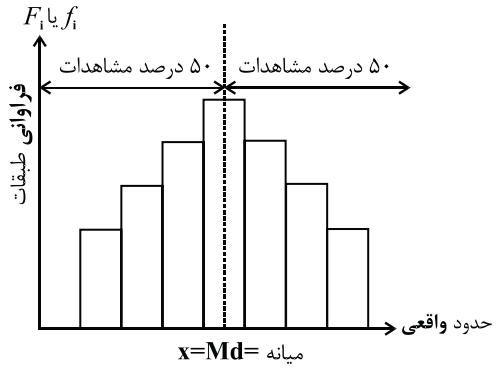
۷. *از نظر هندسی، میانه، خطی عمودی به معادله ..... است که نمودار ..... را به ..... تقسیم می‌کند.*

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۱

۷. از نظر هندسی (نموداری)، میانه، خطی عمودی به معادله $X = Md$ می‌باشد که نمودار **هیستوگرام (بافت نگار یا مستطیلی) را به دو سطح مساوی تقسیم می‌کند.**

چون:

همان‌طور که می‌دانیم میانه، داده‌ای است که در **وسط** داده‌های مرتب شده (صعودی) قرار داره، یعنی نصف مشاهدات قبل از آن و نصف دیگر مشاهدات بعد از آن قرار دارند:



نکته مهم:

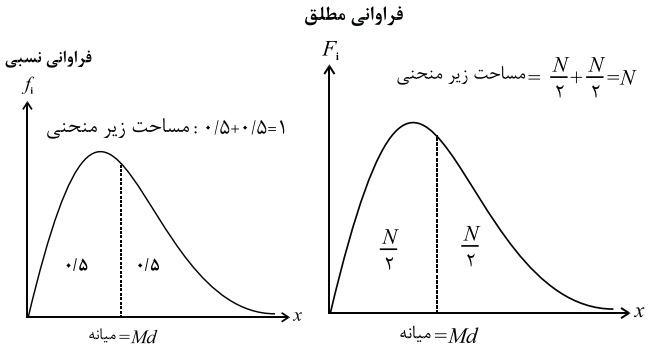
۱) اگر محور عمودی بیانگر فراوانی **مطلق** (F_i) با شد، آنگاه مساحت زیر منحنی توزیع، برابر با **جمع فراوانی‌های مطلق** $(\sum F_i = N)$ است.

(ادامه تو فیش بعد)

۲) و اگر محور عمودی نشـــان دهندهٔ فراوانی **نســبی** (f_i) باشـــد، آنگاه مســـاحت زیر منحنی توزیع برابر با **جمع فراوانی‌های نســبی** ($\sum f_i = ۱$) است.

توجه: علاوه بر ۷ ویژگی (خاصیت) میانه که در این فیش و فیشهای قبل بهش اشـــاره کردیم، میانه یک خاصـــیت بســـیار مهم هم داره که بخاطر اهمیت خیلی زیادش، اونو جداگونه تو فیش‌های بعد بررسی می‌کنیم.

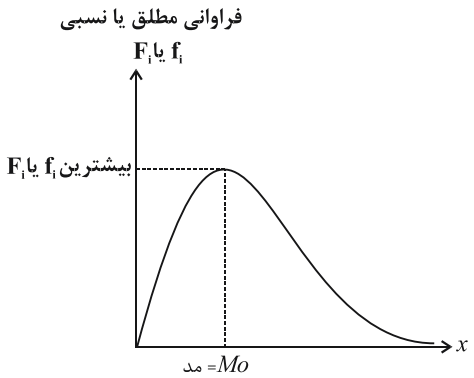
بـا توجـه بـه اینکـه میانـه، مقـداری از متغیـر مربوطـه اسـت کـه روی محـور X هـا قـرار داره و داده‌هـا را بـه **دو قســمت مســاوی** تقســیم مـی‌کنـه، پـس از نظـر هندسـی، **میانـه نقطـه ای روی محـور X هاسـت** که **مساحت توزیع را به دو نیم تقسیم می‌کند**:



(ادامه تو پشت فیش)

مقایسه میانه و مد (از لحاظ نموداری):

مد، مقداری روی محور X ها بود که دارای بیشـــترین فراوانی (مطلق یا نسبی) بود، یعنی منحنی توزیع به ازای مد، دارای ***بیشترین ارتفاع*** بود:



(بسیار مهم):

خاصیت مهم میانه (Md)

مهمترین خاصـــیت میانه به یکی از حالت‌های زیر می‌تواند بیان شود:

۱) مجموع **قدرمطلق** انحرافات (تفاضلات) داده ها از **میانه**، ..... است. از این خاصیت **میانه** در چه مواردی استفاده می‌شود؟

---

۲) مجموع **قدرمطلق** انحرافات (تفاضلات) داده‌ها از **میانه**، از مجموع **قدرمطلق** انحرافات (تفاضلات) داده‌ها نسبت به هر نقطه یا عدد دلخواه دیگری ..... است.

۳) هرگاه مجموع **قدرمطلق** انحرافات (تفاضـــلات) داده‌ها از نقطه دلخواهی (مانند C) ..... باشد، آنگاه آن نقطه، **میانه** است و بالعکس.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۳۲

آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار، ص ۹۵

خاصیت مهم میانه:

مهمترین خاصیت (ویژگی) میانه را می‌توان به یکی از حالات زیر بیان کرد:

ا) مجموع **قدرمطلق** انحرافات (تفاضلات) داده‌ها از میانه، همیشه **حداقل** است:

$$\underbrace{\sum |\ |}_{\text{قدرمطلق}} \quad \underbrace{x_i - Md}_{\text{انحرافات}}$$

$$\sum_{i=1}^{k} |x_i - Md| = \min$$

توجه: اگر داده‌ها دارای **فراوانی** (**مطلق** F_i **یا نسبی** f_i) باشند خاصیت فوق به صورت زیر تبدیل می‌شود:

$$\begin{cases} \sum F_i |x_i - Md| = \min & \text{در صورت داشتن فراوانی } \textbf{مطلق}: \\ \sum f_i |x_i - Md| = \min & \text{در صورت داشتن فراوانی } \textbf{نسبی}: \end{cases}$$

نکته مهم: از این خاصیت مهم میانه برای طرح‌ریزی خطوط راه آهن برقی شهرها، اتوبوس، مترو و در تعیین ایستگاه‌ها، محل پمپ‌های بنزین، انبارهای عمومی و ... استفاده می‌شود.

مثال: *فرض کنید می‌خواهیم در یک جاده، پمپ بنزینی را احداث کنیم که چندین شرکت مسافربری از این پمپ بنزین، استفاده می‌کنند.*

(ادامه تو فیش بعد)

برای ما مهم است که بدانیم پمپ بنزین را در چه محلی احداث کنیم، تا در مجموع، کمترین مسافت توسط اتوبوس‌های مسافربری برای سوخت‌گیری پیموده شود.

***مثال۲:** فرض کنید یک دانشگاه دارای چند دانشکده است که هر دانشکده نیز دارای تعدادی دانشجو می‌باشد. می‌خواهیم محلی را جهت احداث سلف سرویس انتخاب کنیم به طوری که مسافت پیموده توسط افراد به حداقل خودش برسد.*

*در هر یک از این ۲ مثال باید از **خاصیت مهم میانه** استفاده کنیم یعنی پمپ بنزین را باید در **میانه** فواصل شرکت‌های مسافربری تا ایستگاه پمپ بنزین احداث کنیم تا مجموع فاصله اتوبوس‌های این شرکت‌ها تا ایستگاه پمپ بنزین حداقل شود.*
*سلف دانشگاه را نیز باید در **میانه** فاصله دانشکده‌های مختلف از سلف بسازیم تا بدین ترتیب مجموع فاصله دانشکده‌های مختلف از سلف حداقل گردد:*

$$\boxed{\sum |x_i - Md| = \min}$$

در این رابطه، عبارت $|x_i - Md|$ نشان دهنده **فاصله** داده‌ها (یعنی فاصلهٔ شرکتهای مسافربری، دانشکده‌های یک دانشگاه و...) از میانه (Md) است.

(ادامه تو پشت فیش)

۲) خاصیت ۱ را می‌توان به صورت زیر نیز بیان نمود:

مجموع **قدرمطلق** انحرافات (تفاضلات) داده‌ها از **میانه** از مجموع **قدرمطلق** انحرافات (تفاضلات) داده‌ها نسبت به هر نقطه یا عدد دلخواه دیگری غیر از میانه: $(a \neq Md)$ **کوچک‌تر** است:

$$\sum |x_i - Md| < \sum |x_i - a|$$

در صورت داشتن فراوانی **مطلق**: $\sum F_i |x_i - Md| < \sum F_i |x_i - a|$

در صورت داشتن فراوانی **نسبی**: $\sum f_i |x_i - Md| < \sum f_i |x_i - a|$

۳) خاصیت مهم میانه را هم چنین می‌توان به شکل زیر بیان کرد.

هر گاه مجموع **قدرمطلق** انحرافات (تفاضلات) داده ها از نقطه دلخواهی (مانند C) حداقل باشد، آنگاه نتیجه می گیریم که آن نقطه، **میانه** است و بالعکس، یعنی اگر **میانه** مشاهدات را به عنوان یک نقطه دلخواه $(Md = C)$ در نظر بگیریم، آنگاه مجموع **قدرمطلق** انحرافات داده‌ها از این نقطه دلخواه $(Md:$ **میانه**$)$ **حداقل** خواهد بود:

$$\sum |x_i - C| = \min \xleftrightarrow{\text{دو طرفه}} C : Md$$

(ادامه تو فیش بعد)

با داشتن فراوانی **مطلق** : $\sum F_i |x_i - C| = \min \rightleftarrows C = Md$

با داشتن فراوانی **نسبی** : $\sum f_i |x_i - C| = \min \rightleftarrows C = Md$

نکته بسیار مهم در تست‌ها: باید توجه داشت که این خاصیت مهم میانه، همواره باید به صورت **قدرمطلق** بیان شود، بنابراین در تست‌ها برای به اشتباه انداختن داوطلبان، این خاصیت را گاهی بدون قدرمطلق: $\sum (x_i - Md)$ و یا به صورت **مجذور**: $\sum (x_i - Md)^2$ نشان می‌دهند که ما باید حواسمون جمع باشه که هیچکدوم از اینا نشون دهندهٔ خاصیت مهم میانه نیستن:

غلط است زیرا از **قدرمطلق** استفاده نکرده است $\sum (x_i - Md) = \min \rightarrow$

غلط است زیرا از **توان ۲** استفاده کرده است $\sum (x_i - Md)^2 = \min \rightarrow$ و از قدرمطلق استفاده نکرده.

غلطه، زیرا قدرمطلق ما نباید توان داشته باشد $\sum |x_i - Md|^2 = \min \rightarrow$

خب، الان می‌خواهیم این خاصیت مهم میانه را با یک مثال عددی نشون بدیم:

مثال: *اگر مشاهدات ما شامل ۱۵، ۱۰ و ۵ باشند، آنگاه:* $Md = ۱۰$: **میانه**

حالا می‌خواهیم به مقایسهٔ مقادیر $\sum |x_i - Md|$ و $\sum |x_i - a|$ بپردازیم (**توجه**: $a \neq Md$ است):

(ادامه تو پشت فیش)

x_i صعودی	$a \neq Md$ $\lvert x_i - ۵ \rvert$	$a \neq Md$ $\lvert x_i - ۷/۵ \rvert$	Md $\lvert x_i - ۱۰ \rvert$	$a \neq Md$ $\lvert x_i - ۱۲/۵ \rvert$
۵	$\lvert ۵-۵ \rvert = ۰$	$\lvert ۵-۷/۵ \rvert = ۲/۵$	$\lvert ۵-۱۰ \rvert = ۵$	$\lvert ۵-۱۲/۵ \rvert = ۷/۵$
$\boxed{۱۰}$ = **میانه**	$\lvert ۱۰-۵ \rvert = ۵$	$\lvert ۱۰-۷/۵ \rvert = ۲/۵$	$\lvert ۱۰-۱۰ \rvert = ۰$	$\lvert ۱۰-۱۲/۵ \rvert = ۲/۵$
۱۵	$\lvert ۱۵-۵ \rvert = ۱۰$	$\lvert ۱۵-۷/۵ \rvert = ۷/۵$	$\lvert ۱۵-۱۰ \rvert = ۵$	$\lvert ۱۵-۱۲/۵ \rvert = ۲/۵$
جمع: $\sum$	$\sum \lvert x_i - ۵ \rvert =$ ۱۵	$\sum \lvert x_i - ۷/۵ \rvert =$ ۱۲/۵	$\sum \lvert x_i - ۱۰ \rvert = \boxed{۱۰}$ min	$\sum \lvert x_i - ۱۲/۵ \rvert$ $= ۱۲/۵$

با دقت در جدول بالا می‌بینیم که مجموع قدرمطلق تفاضلات (انحرافات) داده‌ها از میانه $(Md = ۱۰)$ حداقل است:

$$\sum \lvert x_i - Md \rvert = \sum \lvert x_i - ۱۰ \rvert = ۱۰ \quad (\min = \text{حداقل})$$

و همچنین می‌تونیم بگیم که:

مجموع **قدرمطلق** تفاضلات داده ها از **میانه** $(Md = ۱۰)$ از مجموع **قدرمطلق** تفاضلات داده‌ها از هر عدد دیگه‌ای غیر میانه (مثل $a = ۵ \ , \ ۷/۵ \ , \ ۱۲/۵$) **کمتره**:

$$\sum \lvert x_i - Md \rvert < \sum \lvert x_i - a \rvert$$

$$\Downarrow$$

$$\begin{cases} \sum \lvert x_i - ۱۰ \rvert = \boxed{۱۰} < \sum \lvert x_i - ۵ \rvert = \boxed{۱۵} \\ \sum \lvert x_i - ۱۰ \rvert = \boxed{۱۰} < \sum \lvert x_i - ۷/۵ \rvert = \boxed{۱۲/۵} \\ \sum \lvert x_i - ۱۰ \rvert = \boxed{۱۰} < \sum \lvert x_i - ۱۲/۵ \rvert = \boxed{۱۲/۵} \end{cases}$$

(حسابداری ۷۹)

خاصیت مهم میانه (Median) *برای داده‌های آماری عبارت است از مجموع:*

۱) انحرافات از میانه صفر است.

۲) مجذور انحرافات از میانه حداقل است.

۳) قدرمطلق انحرافات از میانه صفر است.

۴) قدرمطلق انحرافات از میانه، از مجموع قدرمطلق انحرافات از هر عدد دیگری کمتر است.

توجه: علت نادرست بودن سایر گزینه‌ها را نیز بیان کنید.

پاسخ: گزینه ۴

خاصیت مهم میانه: مجموع **قدرمطلق** انحرافات داده‌ها (x_i ها) از میانه (me) همیشه حداقل (مینیمم) است: $\sum|x_i - me| = \min$

و یا میشه گفت:

مجموع **قدرمطلق** انحرافات داده ها از میانه از مجموع **قدرمطلق** انحرافات داده‌ها از هر عدد دیگری غیر میانه ($a \neq Md$) **کمتر** است:

$$\sum|x_i - me| < \sum|x_i - a|$$

اما سایر گزینه‌ها غلط‌اند، چون:

گزینه ۱: مجموع انحرافات از میانگین (نه میانه) صفر است.

$$\boxed{\sum|x_i - \mu| = ۰}$$

گزینه ۲: مجموع محذور انحرافات از میانگین (نه میانه) حداقل است.

$$\boxed{\sum(x_i - \mu)^۲ = \min}$$

گزینه ۳: قدرمطلق انحرافات از میانه، حداقل است (نه صفر).

$$\boxed{\sum|x_i - me| = \min}$$

مهم:

(مدیریت ۸۴)

در ۵۰ داده آماری، میانه $Md = ۱۲$, $\sum_{i=۰}^{۵۰} x_i = ۵۵۰$ است، اگر $A = \sum_{i=۱}^{۵۰} |x_i - ۱۲|$ و $B = \sum_{i=۱}^{۵۰} |x_i - ۱۱|$ باشد، آنگاه:

۱) $A < B$

۲) $A = B$

۳) $A > B$

۴) $A = B - ۱$

گزینه ()

خاصیت مهم میانه:

a: هر عدد دلخواه **غیر میانه** $\Leftarrow \sum|x_i - Md| < \sum|x_i - a|$

در این سؤال:

$$\begin{cases} Md = ۱۲ \\ \sum x_i = ۵۵۰ \Rightarrow \mathbf{m} = \dfrac{\sum \mathbf{x_i}}{\mathbf{N}} = \dfrac{۵۵۰}{۵۰} = \dfrac{۵۵}{۵} \Rightarrow \boxed{\mu = ۱۱} \\ N = ۵۰ \end{cases}$$

بنابراین می‌تونیم بگیم:

$$\sum|x_i - \underset{\downarrow}{Md}| < \sum|x_i - \underset{\downarrow}{a}|$$

$$\underbrace{\sum|x_i \quad ۱۲|}_{A} < \underbrace{\sum|x_i - ۱۱|}_{B}$$

مهم:

فرض کنید بین دو شهر که فاصله آنها ۱۰۰ کیلومتر است، ۷ گاراژ وجود داشته باشد، در جدول زیر تعداد اتومبیل‌های هر گاراژ داده شده و فواصل گاراژها نیز مشخص شده است:

فواصل گاراژ	۰-۲۰	۲۰-۴۰	۴۰-۵۰	۵۰-۸۰	۸۰-۱۰۰
تعداد اتومبیل	۵۰	۱۰۰	۲۰	۴۰	۱۲۰

می‌خواهیم یک پمپ بنزین در این جاده احداث کنیم.

محلی را تعیین کنید که جمع کل مسافت‌های پیموده شده برای بنزین گیری توسط ماشین‌ها، حداقل باشد؟

۱) در فاصله ۴۷/۵ کیلومتری

۲) در فاصله ۴۵ کیلومتری

۳) در فاصله ۴۹/۵ کیلومتری

۴) در فاصله ۴۲/۵ کیلومتری

آمار و کاربرد آن در مدیریت ۱، مسعود نیکوکار، ص ۹۵، مثال ۳۳

پاسخ: گزینه «۱»

فرض کنید نقطه a *محل احداث پمپ بنزین باشد و* x_i *محل گاراژ* i *اُم. خب، پس الان می‌تونیم بگیم مسافتی که هر ماشین از گاراژِ* i *اُم تا رسیدن به پمپ بنزین باید طی کنه، برابره با:* $|x_i - a|$.

توجه: در ریاضیات، برای نشون دادن فاصله (مسافت) بین ۲ نقطه، از مفهوم **قدرمطلق** استفاده می‌کنیم، بنابراین فاصله گاراژ iاُم تا محل احداث پمپ بنزین (یعنی نقطه a) برابر میشه با:

$|x_i - a|$ = محل پمپ بنزین − محل گاراژ iاُم	: **فاصله** گاراژ iاُم تا پمپ بنزین

خب، مسئله از ما خواسته محل پمپ بنزین (a) رو طوری مشخص کنیم تا مجموع فاصله گاراژهای مختلف تا پمپ بنزین (نقطه a) حداقل بشه، یعنی:

حداقل =	پمپ بنزین − محل گاراژ iاُم	$\sum$: مجموع فواصل گاراژها تا پمپ بنزین

که به صورت ریاضی میشه: $\boxed{\sum |x_i - a| = \min}$

خب، با دقت در عبارت **min** و علامت **قدرمطلق** می‌بینیم که این رابطه برامون خیلی آشناست، انگار خیلی وقته می‌شناسیمش. ببینیم یادتون اومد؟ ... درسته، این رابطه در واقع همون **خاصیت مهم میانه** است که به ما میگه: **(ادامه تو فیش بعد)**

اگه مجموع **قدرمطلق** تفاضل داده‌ها (x_i ها) از یه نقطه (مثـل a) حـداقل باشه (min)، اون وقت با جرأت هر چه تمام می‌تونیم بگیم کـه اون نقطـه (یعنی a)، حتماً **میانه** این داده‌هاست:

$$\sum |x_i - a| = \min \quad \Leftrightarrow \quad \text{a} = med : \text{میانه}$$

خب، معما چو حل گشت آسان شود، یعنی الان فکر کنم همه مـون دیگـه فهمیدیم که منظور سؤال از اینکه گفته محل نقطه a (محل احداث پمـپ بنزین) رو بدست بیارین، در واقع اینه که **میانهٔ** x_i ها، یعنی **میانه** فاصـله گاراژها از همدیگه رو بدست بیاریم.

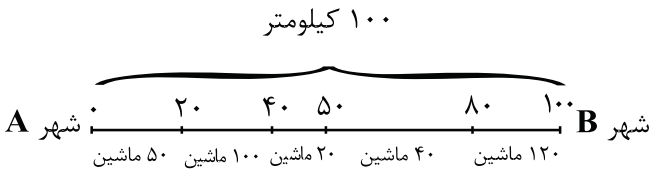
یادآوری:

البته اگه یه کم زرنگی می‌کردیم، زودتر از اینها هم می‌تونستیم بفهمیم که این سؤال مربوط به بحث «**میانه**» است، چون تو فیشای قبل یاد گـرفتیم که یکی از کاربردهای **خاصیت مهم میانه** در زمینه تعیـین احـداث **پمـپ بنزین‌ها**، ایستگاه‌های اتوبوس و مترو و تعیـین محـل احـداث انبارهـا و ... است.

خـب، حـالا کـه فهمیـدیم قضـیه چیـه، بایـد خیلـی سـریع بـریم سـراغ محاسبه میانه؛ امـا قبـل از اینکـه بـه جـدول پشـت فـیش و محاسـباتش نگـاه کنـین، بهتـره اول بـه شـکل رسـم شـده دقـت کنـین تـا مفهـوم سؤال رو بهتر متوجه بشین.

(ادامه تو پشت فیش)

(در این شکل فاصله گاراژها بر روی محور و تعداد ماشین‌های هر گاراژ در زیر محور نوشته شده).



دقت کنین که: *فواصل گاراژها، در واقع طبقات ما هستن و تعداد ماشین‌های هر گاراژ، فراوانی هر یک از این طبقات هستن، بنابراین جدول ما به صورت زیر درمیاد:*

گام ۱) محاسبه فراوانی تجمعی طبقات (در جدول زیر):

طبقه میانه دار

CL	۰-۲۰	۲۰-۴۰	۴۰ – ۵۰	۵۰-۸۰	۸۰-۱۰۰	
F_i	۵۰	۱۰۰	۲۰	۴۰	۱۲۰	$N = ۳۳۰$
F_{C_i}	۵۰	۱۵۰	۱۷۰	۲۱۰	$۳۳۰ = N$	

گام ۲) یافتن محل میانه:

$$\boxed{\text{میانه} : C_{Md} = \frac{N}{۲}} \Rightarrow \frac{۳۳۰}{۲} = ۱۶۵\text{اُمی}$$

(ادامه تو فیش بعد)

اولین طبقه‌ای که فراوانی تجمعی آن بزرگتر مساوی ۱۶۵ است، طبقه (۵۰-۴۰) است که طبقه میانه دار است.

پس طبقه سوم، طبقه میانه داره $\xrightarrow{\text{در طبقه سوم}} F_{C_۳} = ۱۷۰ \geq ۱۶۵ \rightarrow$

$$Md = ۴۰ + \frac{۱۶۵ - ۱۵۰}{۲۰} \times ۱۰ \Rightarrow$$

$$Md = ۴۰ + \frac{۱۵}{\cancel{۲۰}} \times \cancel{۱۰} = ۴۰ + \frac{۱۵}{۲} = ۴۰ + ۷/۵ \Rightarrow$$

$$\boxed{Md = ۴۷/۵ \text{ کیلومتر}}$$

یعنی پمپ بنزین باید در فاصله ۴۷/۵ کیلومتری این جاده احداث بشه تا در نتیجه مجموع مسافت‌های طی شدهٔ این اتومبیل‌ها برای بنزین‌گیری، **حداقل** بشه:

$$\sum |x_i - Md| = \min$$

$$\xrightarrow{Md=۴۷/۵} \boxed{\sum |x_i - ۴۷/۵| = \min}$$

امام علی علیه السلام
السَّیِّئُ الخُلُقِ کَثیرُ الطَّبیشِ مُتَغَّصُ العَیشِ
غررالحکم حدیث ۱۶۰۱۴
آدم بد اخلاق بسیار خطا می کند
و زندگی اش تلخ می شود
بیا از این بد اخلاقی حذر کن
بیا تا میتوانی رفح شرکن
اگر خواهی که با یاران نشینی
محبت را به یاران بیشتر کن

مهم:

۱) در توزیعهای نامتقارن، برای نشون دادن میزان تمایل داده‌ها به مرکز توزیع، بهتر است از کدامیک از شاخص‌های مرکزی استفاده کنیم (مد، میانه یا میانگین) ؟ چرا؟

۲) با توجه به سؤال بالا، توضیح دهید برای توصیف توزیع‌های درآمد، از کدام شاخص مرکزی باید استفاده کنیم؟ چرا؟

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۳۲

«کاربرد میانه»

در شرایط زیر استفاده از میانه (Md) به عنوان شاخص مرکزی مناسب‌تر است:

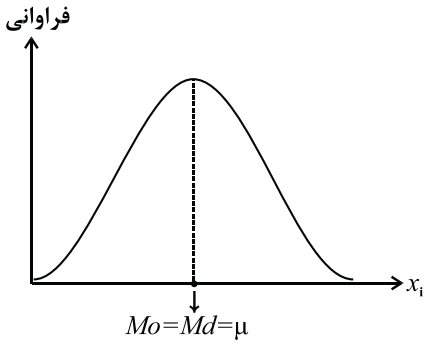
۱. *در توزیعهای نامتقارن:* که تو همین فیش، این مورد رو توضیح می‌دیم.
۲. *در توزیع‌هایی که تعداد اندکی مشاهده غیر طبیعی (بسیار کوچک یا بسیار بزرگ نسبت به سایر مشاهدات) در ابتدا یا انتهای آنها وجود دارد:* که در فیش‌های بعداً این مورد را بررسی می‌کنیم.
۳. *در جدول داده‌های طبقه بندی شده، زمانی که حدود ابتدا یا انتها، باز (نامشخص) باشد:* که این کاربرد میانه را هم در فیش‌های بعد مورد بررسی قرار می‌دهیم.

کاربرد اول: *در توزیعهای نامتقارن، از میانه برای نشون دادن میزان تمایل (گرایش) داده‌ها به مرکز توزیع استفاده می‌شود.*

برای تشریح بهتر حالت فوق، ابتدا بایستی با مفهوم توزیع متقارن و غیرمتقارن آشنا شویم؛ همانطور که در شکل فیش بعد نشان داده شده است، توزیع متقارن حالت **تقارن و قرینگی** دارد، یعنی طرف راست و چپ آن قرینه هم هستند، به همین دلیل به آن، توزیع **متقارن** می‌گویند.

(ادامه تو فیش بعد)

توزیع متقارن به شکل یک **زنگوله** است و در آن تمام شاخص‌های مرکزی یعنی مد، میانه و میانگین بر هم منطبق هستند و درست در وسط یا مرکز توزیع قرار می‌گیرند: $(Mo = Md = \mu)$



اما برخلاف توزیع متقارن که حالت قرینه دارد، توزیع‌های نامتقارن حالت قرینگی ندارند، یعنی طرفین راست و چپ آنها با هم قرینه نیستند، زیرا تو این جور توزیع‌ها، داده‌ها بیشتر در یک سمت توزیع (راست یا چپ) متمرکز شده‌اند.

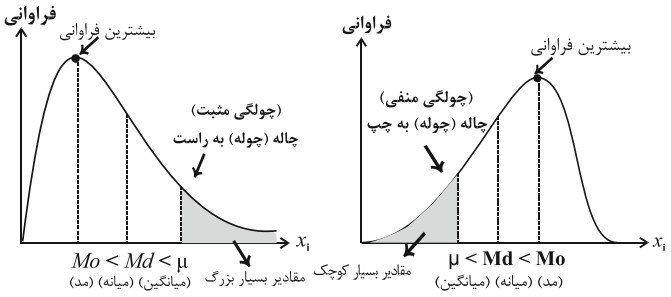
توزیعهای نامتقارن دارای **چولگی** یا **چالگی** هستند، بدین صورت که اگر:

(ادامه تو پشت فیش)

۱) اگر دُم یـا دنبالـه توزیـع در سـمت راسـت قـرار داشـته باشـند (یعنـی اگـه سـمت راسـت توزیـع حالـت **چالـه** داشـته باشـد)، توزیـع **چولـه (چاله)** ***به راست*** بوده و یا دارای ***چولگی مثبت*** است (شکل الف).

۲) اگردُم یا دنباله توزیع در سـمت چـپ قـرار گیـرد (یعنـی اگـه **چالـه** در سمت چپ توزیع ایجـاد شـود)، توزیـع **چولـه (چالـه)** ***به چپ*** بـوده و یا دارای **چولگی منفی** می‌باشد (شکل ب).

در توزیع‌های ***نامتقارن***، بـه دلیـل تجمـع داده‌هـا در یـک طـرف توزیـع (راست یا چـپ)، دیگـر شـاخص‌هـای مرکـزی مـد، میانـه و میـانگین بـر هم منطبق **نیستند**.



(شکل ب) **(شکل الف)**

(ادامه تو فیش بعد)

در توزیع‌های **نامتقارن** (چوله به راست یا به چپ)، چون مقادیر بسیار بزرگ یا بسیار کوچک در ابتدا یا انتهای توزیع وجود داره، بنابراین برای بیان شاخص مرکزی داده‌ها، ***میانه شاخص بهتری نسبت به میانگین و مد است.***

زیرا میانه نسبت به میانگین و نما **کمتر تحت تأثیر مقادیر انتهای توزیع (مقادیر بسیار بزرگ یا بسیار کوچک)** قرار می‌گیرد و **همیشه در وسط توزیع** قرار دارد.

اما میانگین **به شدت** تحت تأثیر مقادیر **فرین (غیر طبیعی: بسیار بزرگ یا بسیار کوچک)** قرار می‌گیرد، بدین معنی که میانگین به سمت مقادیر بسیار بزرگ یا بسیار کوچک گرایش پیدا می‌کند.

این نکته را می‌توان در شکلهای (الف) و (ب) مشاهده کرد:

در شکل (الف) که مقادیر فرین (بسیار بزرگ) در سمت راست توزیع قرار گرفته‌اند، میانگین نیز تحت تأثیر این مقادیر قرار گرفته و به سمت آنها (به سمت راست) نزدیک شده.

و تو شکل (ب) که مقادیر فرین (بسیار کوچک) در سمت چپ توزیع قرار گرفته‌اند، میانگین نیز تحت تأثیر این مقادیر فرین قرار گرفته و به سمت آنها (به سمت چپ) نزدیک شده.

(ادامه تو پشت فیش)

بنابراین در توزیع‌های نامتقارن (شکلهای الـف و ب)، میـانگین نمـی‌توانـد بخوبی مرکز (وسط) توزیع داده‌ها را نشان دهد، زیرا بجای اینکه در مرکـز (وسط) قرار گیرد، در سمت راست یا چپ توزیع قرار می‌گیرد؛

بنابراین در توزیع‌های **نامتقارن** بایستی از شاخص مرکـزی **میانه** بجـای **میانگین و مد** استفاده کنیم، چون میانه **همیشه** در **وسـط** توزیـع قـرار می‌گیرد.

توجه: در توزیع‌های **نامتقارن** از مد نیز نباید استفاده کرد، زیرا همان‌طور که در شکلهای (الف) و (ب) می‌بینـیم، در توزیـع‌هـای نامتقـارن بـرخلاف توزیع‌های متقارن، مد در وسط (مرکز) توزیع داده‌ها قرار نمی‌گیـرد، بلکـه در سمت راست یا چپ توزیع قرار می‌گیرد، پس تو این جور مواقع (یعنـی تو توزیع‌های نامتقارن)، مد شاخص مرکزی خوبی نیست، پس بجاش **باید از میانه** استفاده کنیم. چون میانه همیشه (چه در توزیـع‌هـای متقـارن و چه در توزیع‌های نامتقارن) در **وسط** یا **مرکز** توزیع قرار می‌گیره.

پاسخ سؤال (۲)

نکته مهم: توزیع‌های درآمد، اغلب دارای مقادیر بزرگ فرین (تعداد کمـی درآمد خیلی بالا) هستند، یعنی دارای چولگی **مثبت** (چـولگی بـه **راسـت**) هستند، **(ادامه تو فیش بعد)**

بنابراین این توزیع‌ها متقارن نیستند بلکه نامتقارن اند، پس برای نشان دادن مرکز توزیع‌های درآمد، **میانه شاخص مرکزی بهتری نسبت به مد و میانگین** است، زیرا میانه نسبت به مد و میانگین **کمتر** تحت تأثیر مقادیر فرین (درآمدهای خیلی بالا) قرار می‌گیرد و در نتیجه مرکز توزیع درآمد را بهتر نشان می‌دهد.

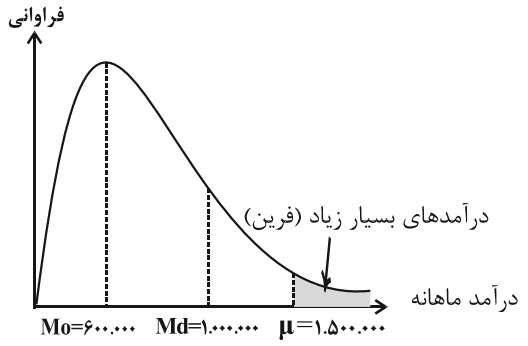
توجه: *در توزیع‌های درآمد، معمولاً عده زیادی از افراد جامعه، درآمد کمی دارند (که مد جامعه را تشکیل می‌دهند، زیرا بیشترین فراوانی را دارند)، ولی عده کمی از افراد جامعه، درآمدهای بسیار بالایی دارند که همان مقادیر فرین هستند که این مقادیر، **به شدت** میانگین را تحت تأثیر خود قرار میدن و اونو به سمت خودشون (به سمت راست) می‌کشن، بنابراین در این حالت، میانگین دیگر در وسط (مرکز) توزیع درآمد قرار نمی‌گیرد، بلکه به سمت راست کشیده می‌شود؛*

بنابراین تو این شرایط (توزیع نامتقارن، مثل توزیع درآمد)، دیگه میانگین شاخص مرکزی مناسبی نیست، چون ما از یه شاخص **مرکزی** *انتظار داریم که* **مرکز** *داده‌ها رو به مان شون بده، اما در توزیع‌های نامتقارن (مثل توزیع درآمد)، میانگین دیگه نمی‌تونه بخوبی مرکز داده‌ها رو به ما نشون بده، چون* **مقدارش به سمت مقادیر انتهایی توزیع متمایل میشه (و نه مقادیر مرکز توزیع).** پس:

نتیجه اخلاقی مهم:

در توزیع‌های **نامتقارن** (مثل توزیع درآمد)، برای نشان دادن مرکز داده‌ها، بهتر است از شاخص‌های مرکزی **میانه** بجای شاخص مرکزی مد یا میانگین استفاده کنیم.

به عنوان مثال به شکل توزیع درآمد در زیر توجه کنید:



در این توزیع، همان‌طور که از شکل پیداست، اکثر افراد جامعه، درآمد کمی در حدود ۶۰۰ هزار تومان در ماه دارند و در نتیجه این گروه، بخاطر فراوانی زیادشان (نسبت به سایر گروه‌های درآمدی)، مُدِ جامعه را تشکیل می‌دهند. و در مقابل، تعداد کمی از افراد جامعه، درآمدهای غیرطبیعی (بسیار بزرگ نسبت به بقیه) دارند؛ **(ادامه تو فیش بعد)**

مثلاً درآمدی حدود ۸ میلیون تومان در ماه دارند و همین امر باعث شده تا میانگین به شدت تحت تأثیر درآمدهای خیلی بالا قرار گیرد و به سمت درآمدهای بزرگ نزدیک شود که این امر با واقعیت تطابق ندارد؛

یعنی میانگین واقعی درآمد افراد، مطمئناً کمتر از μ=1,500,000 تومان در ماه خواهد بود، زیرا اکثر افراد جامعه درآمدی حدود 600,000 تا 1,000,000 تومان در ماه دارند، ولی میانگین، به اشتباه در آمد متوسط μ=1,500,000 تومان در ماه را به ما نشون میده.

پس در این حالت (در توزیعهای **نامتقارن**، مثل: توزیع درآمد) بهتر است از شاخص مرکزی **میانه** بجای **مد** و **میانگین** استفاده کنیم، زیرا همان‌طور که در شکل پیداست، ***میانه همیشه در وسط توزیع قرار می-گیرد*** و در نتیجه در توزیع‌های **نامتقارن**، بهتر می‌تونه وسط یا مرکز توزیع رو به ما نشون بده.

حُسن خُلق:

إن العبد لیدرک بحسن الخلق درجة الصائم القائم

بنده بوسیله خوش خلقی به مقام روزه دار و نماز گزار می رسد.

پیامبر اکرم (ص)، نهج الفصاحه/ حدیث٦٥٤

مهم:

*در توزیع‌هایی که تعداد **اندکی** مشاهده غیرطبیعی (**بسیار کوچک یا بسیار بزرگ** نسبت به سایر مشاهدات) در ابتدا یا انتهای آنها وجود دارد، کدام یک از شاخص‌های مرکزی (میانه یا میانگین) برای تعیین نقطه مرکزی داده‌ها مناسب‌تر است؟ چرا؟*

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۳۳

«دومین کاربرد میانه»

در فیش‌های قبل یکی از کاربردهای میانه را بررسی کردیم بدین صورت که:

در توزیع‌های نامتقارن (چوله به راست یا چپ) بایستی از **میانه** بجای مد و میانگین به عنوان شاخص مرکزی استفاده کرد.

اکنون به یکی دیگر از کاربردهای میانه می‌پردازیم:

در توزیع‌هایی که تعداد اندکی مشاهده غیرطبیعی (خیلی کوچک یا خیلی بزرگ) در ابتدا یا انتهای آنها وجود دارد، **میانه** شاخص مرکزی بهتری نسبت به میانگین است.

برای مثال در N=۱۸ مشاهده زیر:

$\boxed{-۲۰۰}$,۲,۲,۲,۳,۳,۴,۴,۴,۵,۵,۵,۶,۶,۶,۶,$\boxed{۱۰۰}$,$\boxed{۴۰۰}$

همانطور که مشخص است داده‌های ۲۰۰- (در ابتدا) و ۱۰۰ و ۴۰۰ (در انتها) اختلاف زیادی با سایر مشاهدات دارند و تعدادشان نیز اندک است (۳مشاهده)؛ در این شرایط پارامتر مرکزی **میانه** نسبت به میانگین، کمتر تحت تاثیر مقادیر ابتدایی و انتهایی (۲۰۰- و ۱۰۰ و ۴۰۰) قرار می‌گیرد، در نتیجه بهتر می‌تونه **مرکز** داده‌ها رو به ما نشون بده، به همین خاطر می‌تونیم بگیم که در این جور مواقع، **(ادامه تو فیش بعد)**

میانه شاخص مرکزی بهتری نسبت به میانگین خواهد بود.

برای درک بهتر این موضوع به محاسبات زیر توجه کنین:

در مثال فوق:

$$\text{میانگین} = \mu = \frac{-۲۰۰+۲+۲+.....+۱۰۰+۴۰۰}{۱۸} = \frac{۳۶۳}{۱۸} \approx ۲۰$$

یعنی میانگین داده‌های فوق برابر $\mu = ۲۰$ است، ولی میانه برابر $Md = ۴/۵$ است.

یادآوری: چون تعداد داده‌ها زوجه $(N = ۱۸)$، پس دو تا داده، در وسط توزیع قرار می‌گیرن که میانه برابر با میانگین این دو دادهٔ وسط خواهد بود:

محل میانه : $C_{Md} = \frac{N}{۲} + \frac{۱}{۲} = \frac{۱۸}{۲} + \frac{۱}{۲} = ۹/۵$ اُمی $\rightarrow$ (بین داده ۹اُم و ۱۰اُم)

میانه : $Md =$ میانگین دو دادهٔ وسط $= \frac{X_۹ + X_{۱۰}}{۲} = \frac{۴+۵}{۲} \Rightarrow \boxed{Md = ۴/۵}$

با توجه به مثال بالا می‌بینم که مقدار میانه $(Md = ۴/۵)$ نسبت به میانگین $(\mu = ۲۰)$، به مرکز داده‌ها بیشتر نزدیکه، پس نتیجه می‌گیریم که در اینجا میانه شاخص مرکزی بهتری نسبت به میانه است.

(ادامه تو پشت فیش)

نکته مهم: البته بعداً خواهیم دید که برای رفـع ایـن عیـب در شـاخص میانگین حسابی، باید از میانگین ***پیراسته*** بجای میانگین **حسابی** استفاده کنیم.

در میانگین پیراسته، داده‌هایی که در ابتدا یا انتهای توزیع قرار دارند و در نتیجه اختلاف زیادی با بقیه داده‌ها دارند، رو حـذف مـی‌کنـیم و فقط میانگین داده‌های باقیمونده رو حساب می‌کنیم.

در نتیجه با این ترفندی که بکار می‌بریم، دیگه مقدار میـانگین تحـت تاثیر این مقادیر ابتدایی و انتهایی توزیع قرار نمی‌گیره.

(مهم):

اگه در جدول داده‌های طبقه‌بندی شده (به صورت فاصله‌ای)، حدود ابتدا و انتها **باز (نامشخص)** *باشند، باید از کدام شاخص مرکزی (میانه یا میانگین) استفاده کنیم؟ چرا؟*

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۳

«سومین کاربرد میانه»

در فیش‌های قبل، دو مورد از کاربردهای میانه را بررسی کردیم:

۱. در توزیع‌های نامتقارن (چوله به راست و چپ)

۲. در توزیع‌هایی که تعداد اندکی مشاهده غیرطبیعی (بسیار کوچک یا بسیار بزرگ) در ابتدا یا انتهای آنها وجود دارد.

خب، حالا نوبت سومین کاربرد میانه است:

در جدول داده‌های طبقه‌بندی شدهٔ فاصله‌ای، زمانی که **حدود ابتدا یا انتها باز (نامشخص) باشند**، بجای شاخص میانگین باید از شاخص مرکزی **میانه** استفاده کنیم.

مثال: *اگر مشاهدات ما بصورت زیر باشد:*

۲۰۰-,۲,۲,۲,۳,۳,۴,۴,۴,۵,۵,۵,۶,۶,۶,۶,۱۰۰,۴۰۰

می‌توان آنها را بصورت زیر طبقه‌بندی نمود:

CL	<۲ (کمتر از ۲)	۲-۵	۵-۸	≥ ۸	
فراوانی	۱	۸	۷	۲	$\sum F_i = N = ۱۸$

(ادامه تو فیش بعد)

توضیح: چون در مشاهدات فوق، تنها یک مشاهده قبل از x=۲ وجود دارد (۲۰۰-) که فاصله زیادی با طبقه دوم (۵-۲) دارد، پس ما دیگه نمی-تونیم اونو مثل طبقات دوم و سوم بصورت (حد بالا – حد پایین) نشون بدیم، چون در این‌صورت بایستی تعداد زیادی طبقه با فاصله I=۳ ایجاد می‌کنیم تا به داده ۲۰۰- برسیم.

در مورد داده‌های ۱۰۰ و ۴۰۰ نیز همچنین قضیه صادق است، یعنی اگر بخواهیم ۱۰۰ و ۴۰۰ طبقاتی رو در جدول مشخص کنیم که بصورت (حدبالا – حد پایین) باشن، باید تعداد زیادی طبقه خالی (یعنی طبقاتی با فراوانی صفر ایجاد کنیم) تا بالاخره به ۱۰۰ و ۴۰۰ برسیم که این کار اصلاً کار قشنگی نیست.

پس برای اینکه جدول ما مثل یه تریلی دراز و بی‌قواره نشه، میائیم حد پایین طبقه اول و نیز حد بالای طبقهٔ آخر را **باز** می‌داریم. تا اعداد ۲۰۰- ، ۱۰۰ و ۴۰۰ تو این طبقات **باز** قرار بگیرن.

نکته مهم: تو فیشای بعد یاد می‌گیریم که در داده‌های نوع **سوم** که بصورت **فاصله‌ای** طبقه‌بندی می‌شوند (مثل جدول بالا)، برای محاسبه میانگین از رابطهٔ $\mu = \frac{\sum x_i}{N}$ ، نیاز داریم تا مرکز طبقات x_i ها را از رابطهٔ $x_i = \frac{L_i + U_i}{۲}$ بدست آوریم تا بتوانیم میانگین را محاسبه کنیم.

(ادامه تو پشت فیش)

ولی از اون جایی که تو طبقات باز، حد پایین یا بالا نامشخص است، پس برای این طبقات، نمی‌تونیم مرکز طبقه را از رابطه زیر بدست بیاریم:

$$\text{مرکز طبقه} = \frac{\text{حد پایین طبقه} + \text{حد بالای طبقه}}{۲} \Rightarrow X_i = \frac{L_i + U_i}{۲}$$

در نتیجه تو این حالت (طبقات باز)، «**میانه**» تنها شاخص مرکزی مناسب است، چون برای محاسبه میانه، دیگه نیازی به یافتن مرکز طبقات نداریم و تنها کافی است فراوانی (مطلق یا نسبی) طبقات را بدانیم تا بتوانیم فراوانی تجمعی را از روی آن محاسبه کنیم و سپس محل میانه (طبقه میانه‌دار) را مشخص می‌کنیم (اولین دسته‌ای که فراوانی تجمعی آن از $\frac{N}{۲}$ بیشتر باشد طبقه میانه دار است) و بعد از آن هم مقدار میانه را از رابطه زیر بدست بیاریم:

$$\textbf{میانه} := Md = \text{حد پایین واقعی طبقه میانه دار} + \frac{\frac{N}{۲} - F_{C_{i-۱}}}{F_i} \times \text{طول (فاصله طبقات)}$$

نتیجه: برای محاسبه میانه، هیچ نیازی به دونستن مرکز طبقات نداریم، پس در جداول دارای طبقات **باز** هم (که امکان یافتن مرکز طبقات باز وجود ندارد)، هم می‌تونیم **میانه** رو براحتی آب خوردن حساب کنیم.

«**چندک‌ها**: Quantiles»

۱) **چندک‌ها** چند نوعند؟ و هر کدوم از اونا جامعه رو به چند قسمت تقسیم می‌کنن؟

۲) انواع چارک‌ها رو نام ببرین و بگین هر کدوم از اونا چه مفهومی دارن؟

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳٤

آمار کاربردی ۱، عالِم تبریز و احمد هژبر، ص ۵۰

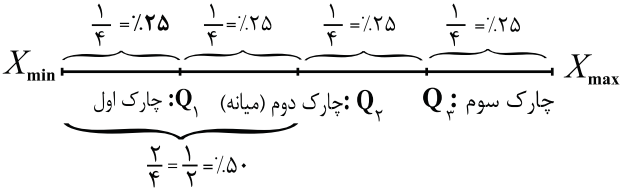
«**چندک‌ها** (Quantile): [**چارک، دهک و صدک**]»

۱) چندک‌ها مقادیری از مشاهدات هستند که دامنهٔ تغییرات جامعه آماری (R) را که بصورت ***صعودی*** مرتب شده، به فواصل **مساوی** با نسبت‌های $\frac{۱}{۴}$ (***چارک***)، $\frac{۱}{۱۰}$ (***دهک***) و $\frac{۱}{۱۰۰}$ (***صدک***) تقسیم می‌کنن، به طوری که فراوانی در هر یک از این فواصل مساوی، درصد معینی از فراوانی کل را تشکیل می‌دهد.

۲) قبلاً یاد گرفتیم که میانه (به عنوان یک شاخص مرکزی) عـددی اسـت که مشاهدات جامعه را به ***دو قسمت مساوی*** تقسیم می‌کند.
خب، حالا اگه تعداد داده‌های ما خیلی زیـاد باشـه (مـثلاً $N > ۳۰$) اون وقت می‌تونیم مفهوم میانه را تعمیم بدیم؛ یعنی می‌تونیم جامعه آمـاری را به ***چهار قسمت مساوی*** تقسیم کنیم که به این نقاطِ تقسیم داده‌هـا بـه ۴ قسمت، اصطلاحاً **چارک** می‌گیم.

چارک (Quartile = Q): اگر دامنه داده‌های جامعه آماری را به ***چهار قسمت مساوی*** تقسیم کنیم، به طوری که هر یک از این قسمتها ۲۵٪ یا $\frac{۱}{۴۰۰}$ کل فراوانی‌ها را در برگیرند، اون وقت چارکهای **اول** $(Q_۱)$ **و دوم** $(Q_۲)$ **و سوم** $(Q_۳)$ بوجود می‌آیند. **(ادامه تو فیش بعد)**

به عبارت بهتر بعد از مرتب کردن داده‌ها به صورت صعودی (غیرنزولی) داریم:



با توجه به نمودار فوق می‌توان نتیجه گرفت که:

$Q_۱ = \frac{۱}{۴}$ **(چارک اول)**

مقداری که $\frac{1}{4} = 25\%$ مشاهدات کوچکتر یا مساوی ($\leq$) آن هستند. یا

مقداری که $\frac{3}{4} = 75\%$ مشاهدات بزرگتر ($>$) از آن هستند.

$Q_۲ = \frac{۲}{۴}$ **(چارک دوم = میانه)**

مقداری که $\frac{2}{4} = 50\%$ مشاهدات کوچکتر یا مساوی ($\leq$) آن هستند. یا

مقداری که $\frac{2}{4} = 50\%$ مشاهدات بزرگتر ($>$) از آن هستند.

$Q_۳ = \frac{۳}{۴}$ **(چارک سوم)**

مقداری که $\frac{3}{4} = 75\%$ مشاهدات کوچکتر یا مساوی ($\leq$) آن هستند. یا

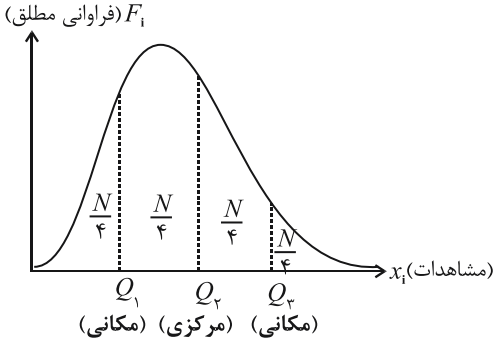
مقداری که $\frac{1}{4} = 25\%$ مشاهدات بزرگتر ($>$) از آن هستند.

(ادامه تو پشت فیش)

نکته: با دقت در نمودار زیر می‌بینیم که $Q_۱$ (چارک اول) مکانی از محور xها است که در موقعیت $\frac{۱}{۴}$ از توزیع قرار داره، یعنی **مکانی** که $\frac{N}{۴}$ از داده‌ها **در پشت سرش** و $\frac{۳N}{۴}$ داده‌ها **در جلوش** قرار دارن (یعنی چارک اول لزوماً **مرکز** توزیع را نشان نمی‌دهد)، به همین دلیل به این شاخص، شاخص **مکانی** می‌گویند (***نه شاخص مرکزی***).

این موضوع در مورد $Q_۳$ (چارک سوم) هم صدق می‌کند، زیرا $Q_۳$ **مکانی** از توزیع است که $\frac{۳N}{۴}$ داده‌ها **در پشت سرش** و $\frac{N}{۴}$ داده‌ها هم **در جلوش** قرار دارند.

ولی $Q_۲$ (چارک دوم = میانه)، یک شاخص **مرکزی** است، زیرا $Q_۲$ بر خلاف $Q_۱$ و در ***وسط و مرکز*** توزیع داده‌ها قرار می‌گیره:



مهم:

(اقتصاد ۷۷)

چارک سوم حقوق در یک سازمان، ۶۵ هزار تومان است. یعنی سه چهارم کارکنان ..... تومان حقوق می‌گیرند.

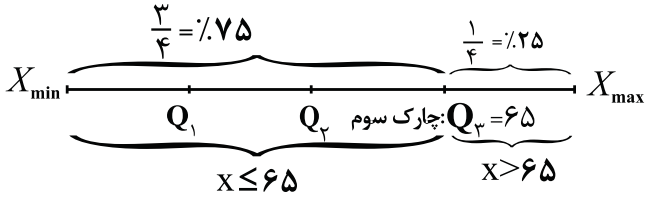
۱) تا ۶۵

۲) ۶۵

۳) بیشتر از ۶۵

۴) ۶۵ هزار تومان و بقیه کارکنان کمتر از ۶۵

پاسخ: گزینه ۱



پس در مورد چارک $(Q_۳ = ۶۵)$ جملات زیر درست اند:

٪۷۵ = $\frac{۳}{۴}$ کارکنان { تا ٦٥ (≤ ۶۵) ، کمتر یا مساوی ٦٥ (≤ ۶۵) ، حداکثر ٦٥ (≤ ۶۵) ، ٦٥ یا کمتر (≤ ۶۵) } حقوق می‌گیرند.

و می‌تونیم این طور هم بگیم که:

٪۲۵ = $\frac{۱}{۴}$ کارکنان بیشتر از ۶۵ (> ۶۵) هزار تومان حقوق می‌گیرند.

مهم:

(اقتصاد ۸۳)

در یک بررسی آماری در خصوص نمرات هوش دانش‌آموزان یک منطقهٔ آموزش و پرورش، کمیت‌های زیر بدست آمده است:

$Q_1 = ۸۰$ (چارک اول)	$\bar{X}=۱۲۰$ (میانگین حسابی)
$Q_۳ = ۱۴۰$ (چارک سوم)	$Md = ۱۱۰$ (میانه)
$S = ۲۰$ (انحراف معیار)	$Mo = ۱۰۷$ (نما)

کدام یک از موارد زیر درست است؟

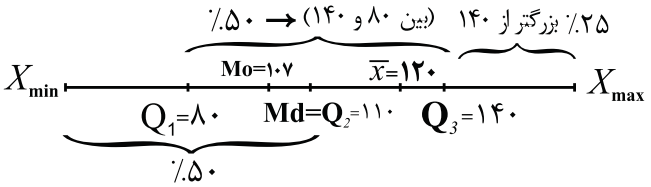
۱) نمرات نیمی از دانش‌آموزان بین ۸۰ تا ۱۴۰ می‌باشد.

۲) نمرات نیمی از دانش‌آموزان کمتر از ۱۲۰ است.

۳) نمرات ۷۵ درصد از دانش‌آموزان بیشتر از ۱۴۰ می‌باشد.

۴) نمرات بیشتر دانش‌آموزان ۱۴۰ و بیشتر است.

پاسخ: گزینه ۱



با توجه به نمودار فوق:

۱) نمرات نیمی از دانش‌آموزان(۵۰%) بین ۸۰ و ۱۴۰ است (گزینه ۱).

۲) چون $\bar{X}=۱۲۰$ بزرگتر از میانه $Md=۱۱۰$ است و با توجه به اینکه نمرات (۵۰٪) از دانش‌آموزان کمتر مساوی $Md=۱۱۰$ است، پس چون ۱۲۰ از ۱۱۰ بزرگتر است (در سمت راست آن قرار دارد)، می‌توان نتیجه گرفت که نمرات بیش از نیمی از دانش‌آموزان کمتر از ۱۲۰ است (گزینه ۲ غلط است).

۳) نمرات ۲۵٪ از دانش‌آموزان بزرگتر از $Q_۳=۱۴۰$ است (گزینه ۳ غلط است).

۴) نمرات بیشتر دانش‌آموزان (مُد نمرات) برابر $Mo=۱۰۷$ است و با توجه به اینکه $Mo=۱۰۷$ قبل از میانه ($Md=۱۱۰$) قرار دارد، پس یعنی بیشتر دانش‌آموزان (بیش از ۵۰٪ دانش‌آموزان) نمرهٔ ۱۰۷ و بیشتر دارند (گزینه ۴ غلط است).

«**دهک** : Decile»

*منظور از **دهک** چیست؟*

*چند نوع **دهک** داریم؟ مفهوم دهک‌های اول، پنجم و نهم را بیان کنید.*

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۵

آمار کاربردی ۱، عالِم تبریز و احمد هژبر، ص ۵۱

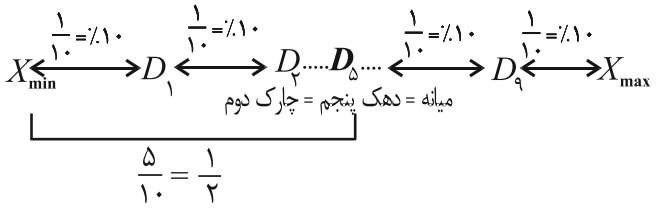
دهک (Decile):

تو فیشای قبلی، با یه شاخص **مکانی** به نام **چارک** آشنا شدیم. خب، حالا می‌خواهیم یه شاخص **مکانی** دیگه رو رونمایی کنیم که اسم مبارکش «**دهک**» است.

شبیه به چارک‌های Q_1, Q_2, Q_3 که توزیع داده‌ها را به **چهار قسمت مساوی** تقسیم می‌کنند، **دهک‌ها** توزیع داده‌ها را به **ده قسمت مساوی** تقسیم می‌کنند که آنها را معمولاً با $D_1, D_2, ..., D_9$ نمایش می‌دهیم.

دهک: *اگر دامنهٔ داده‌های جامعه آماری را به* **۱۰ قسمت مساوی** *تقسیم کنیم، به طوری که هر یک از این قسمت‌ها* **۱۰٪ یا** $\frac{1}{10}$ *از کل فراوانی‌ها را دربرداشته باشند، اون وقت با این کار دهک‌های* **اول** (D_1) **تا نهم** (D_9) *بوجود می‌آیند.*

بعد از مرتب کردن داده‌ها به صورت **صعودی** (غیرنزولی) داریم:



(ادامه تو فیش بعد)

با توجه به نمودار فوق می‌توان نتیجه‌گیری‌های زیر رو داشته باشیم:

$D_{۱} = \frac{۱}{۱۰}$ **(دهک اول)**

مقداری که $\frac{۱}{۱۰} = ٪۱۰$ مشاهدات کوچکتر یا مساوی (≤) آن هستند. یا

مقداری که $\frac{۹}{۱۰} = ٪۹۰$ مشاهدات بزرگتر (<) از آن هستند.

$D_{۵} = \frac{۵}{۱۰}$ **(دهک پنجم = میانه)**

مقداری که $\frac{۵}{۱۰} = ٪۵۰$ مشاهدات کوچکتر یا مساوی (≤) آن هستند. یا

مقداری که $\frac{۵}{۱۰} = ٪۵۰$ مشاهدات بزرگتر (<) از آن هستند.

$D_{۹} = \frac{۹}{۱۰}$ **(دهک نهم)**

مقداری که $\frac{۱}{۱۰} = ٪۱۰$ مشاهدات کوچکتر یا مساوی (≤) آن هستند. یا

مقداری که $\frac{۹}{۱۰} = ٪۹۰$ مشاهدات بزرگتر (<) از آن

(ادامه تو پشت فیش)

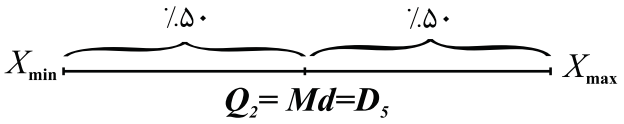
نکته: مشابه نکته‌ای که در مورد چـارک‌هـای اول (Q_1) و سـوم (Q_3) گفتیم، در مورد دهک‌ها نیز می‌تونیم بگیم که:

تنها دهک پنجم (D_5) ***یک شاخص مرکزی*** است، زیرا دقیقاً در **وسط و مرکز** مشاهدات قرار می‌گیرد، ولی بقیه دهک‌ها (D_1 تا D_4 و D_6 تـا D_9) به علت اینکه در مرکز توزیع مشـاهدات قـرار نمـی‌گیرنـد، بنـابراین دیگه شاخص مرکزی نیستند، بلکه تنها **شاخص مکانی توزیع** هستند.

توجه: با دقت در نمودار زیر، می‌توان فهمید که ***دهک پنجم همان میانه یا چارک دوم است***، زیرا دهک پنجم نیز مانند میانه و چارک دوم، مقداری است که $\frac{1}{2} = ۵۰٪$ مشاهدات کوچکتر یا مسـاوی آن و $\frac{1}{2} = ۵۰٪$ بقیـه مشاهدات، بزرگتر از آن هستند:

میانه = چارک دوم = دهک پنجم

$$D_5 = Q_2 = Md$$



دهک چهارم حقوق کارمندان، ۱۲۵ هزار تومان است، این به آن مفهوم است که ..... ۱۲۵ هزار تومان حقوق می‌گیرند.

۱) ۴۰ درصد کارمندان کمتر یا مساوی

۲) ۶۰ درصد کارمندان بیشتر از

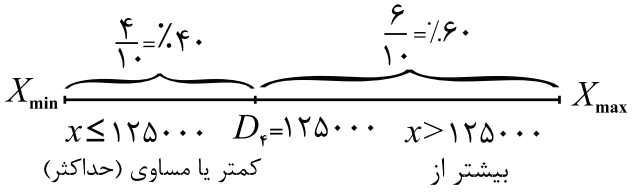
۳) ۴۰ درصد کارمندان

۴) ۶۰ درصد کارمندان حداقل

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۳۶، مثال ۱

پاسخ: گزینه ۱ و ۲ صحیح است.

دهک چهارم $D_۴$، مقداری است که $\frac{۴}{۱۰} = ٪۴۰$ مشاهدات کوچکتر یا مساوی (≤) آن هستند یا $\frac{۶}{۱۰} = ٪۶۰$ مشاهدات بزرگتر (>) از آن هستند.



یعنی:

۴۰٪ کارمندان کمتر یا مساوی ۱۲۵۰۰۰ تومان حقوق می‌گیرند.

۶۰٪ کارمندان بیشتر از ۱۲۵۰۰۰ تومان حقوق می‌گیرند.

«صدک : percentile»

صدک چگونه بوجود می‌آیند؟

مفهوم صدک اول، صدک پنجاهم و صدک نود و نهم چیست؟

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۶

«صدک (percentile :P)»

تو فیشای قبل یاد گرفتیم که شاخص‌های **مکانی** بر ۳ نوع هستن:

۱. چارک‌ها	**۲. دهک‌ها**	**۳. صدک‌ها**

که تا حالا دو مورد اول رو یاد گرفتیم و حالا می‌خواهیم بریم سراغ سومین شاخص **مکانی** که همون **صدک‌ها** هستن:

صدک:اگر دامنه داده‌های جامعه آماری را به **۱۰۰ قسمت مساوی** تقسیم کنیم، به طوری که هر یک از این قسمت‌ها، **یک درصد** $(\frac{۱}{۱۰۰}=۰/۰۱)$ از کل فراوانی‌ها (N) را در برداشته باشد، آنگاه صدک‌های **اول** $(P_۱)$ **تا نود و نهم** $(P_{۹۹})$ بوجود میان.

بعد از مرتب کردن داده‌ها به صورت **صعودی** داریم:

$$X_{min} \xleftrightarrow{\frac{۱}{۱۰۰}=٪۱} P_۱ \xleftrightarrow{\frac{۱}{۱۰۰}=٪۱} P_۲ \ldots \underbrace{P_{۵۰} \ldots}_{\text{میانه = صدک پنجاهم}} \xleftrightarrow{\frac{۱}{۱۰۰}=٪۱} P_{۹۹} \xleftrightarrow{\frac{۱}{۱۰۰}=٪۱} X_{max}$$

$$\frac{۵۰}{۱۰۰}=\frac{۱}{۲}$$

(ادامه تو فیش بعد)

$P_{۱} = \frac{۱}{۱۰}$ **(صدک اول)**

مقداری که $\frac{۱}{۱۰۰} = ۱\%$ مشاهدات کوچکتر یا مساوی (≤) آن هستند. یا

مقداری که $\frac{۹۹}{۱۰۰} = ۹۹\%$ مشاهدات بزرگتر (<) از آن هستند.

$P_{۵۰} = \frac{۵۰}{۱۰۰}$ **(دهک پنجاهم = میانه)**

مقداری که $\frac{۵۰}{۱۰۰} = ۵۰\%$ مشاهدات کوچکتر یا مساوی (≤) آن هستند. یا

مقداری که $\frac{۵۰}{۱۰۰} = ۵۰\%$ مشاهدات بزرگتر (<) از آن هستند.

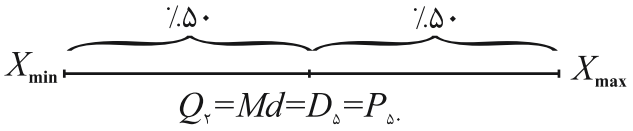
$P_{۹۹} = \frac{۹۹}{۱۰۰}$ **(صدک نود و نهم)**

مقداری که $\frac{۹۹}{۱۰۰} = ۹۹\%$ مشاهدات کوچکتر یا مساوی (≤) آن هستند. یا

مقداری که $\frac{۱}{۱۰۰} = ۱\%$ مشاهدات بزرگتر (<) از آن هستند.

(ادامه تو پشت فیش)

نکته: از بین صدک‌ها، تنها صدک **پنجاهم** $(P_{۵۰})$ است که می‌تواند یک شاخص **مرکزی** محسوب شود، زیرا دقیقاً در **وسط و مرکز** مشاهدات قرار می‌گیرد، ولی سایر صدک‌ها تنها یک شاخص **مکانی** محسوب می‌شوند و نه شاخص **مرکزی**.



(مهم):

«تطبیق چندک‌ها»

۱) از بین انواع چارک‌ها، دهک‌ها و صدک‌ها، کدوم‌شون **منطبق** بر میانه هستند؟

۲) رابطه تبدیل چارک‌ها به صدک چیست؟ (کدوم چارک بر کدوم صدک منطبقه؟)

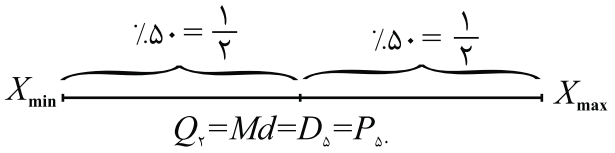
۳) رابطه تبدیل دهک‌ها به صدک‌ها کدامست؟ (کدوم دهک بر کدوم صدک منطبقه؟)

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۳۶

۱) تطبیق چندک‌ها:

شاخص‌های مرکزی: صدک پنجاهم = دهک پنجم = چارک دوم = میانه

$(P_{۵۰})$ $(D_۵)$ $(Q_۲)$ (Md)



۲) تبدیل چارک‌ها به صدک‌ها:

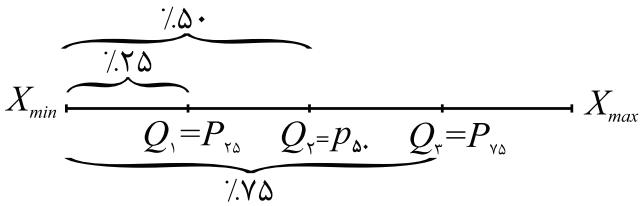
هر چارک در واقع ٪۲۵ = $\frac{۱}{۴}$ داده‌ها را در برمی‌گیرد، پس برای تبدیل چارک‌ها به صدک‌ها می‌توان به طریق زیر عمل کرد:

مثال:

- چارک اول: $Q_۱ \xrightarrow{۱\times۲۵=۲۵}$ صدک بیست و پنجم $(P_{۲۵})$
- چارک دوم: $Q_۲ \xrightarrow{۲\times۲۵=۵۰}$ صدک پنجاهم $(P_{۵۰})$
- چارک سوم: $Q_۳ \xrightarrow{۳\times۲۵=۷۵}$ صدک هفتاد و پنجم $(P_{۷۵})$

(ادامه تو فیش بعد)

بنابراین:

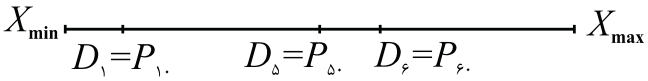


۳) تبدیل دهک‌ها به صدک‌ها:

با توجه به اینکه هر دهک، شامل $\frac{۱}{۱۰}$ = ۱۰٪ داده‌هاست، پس می‌توان گفت:

مثال:

- دهک اول : $D_۱ \xrightarrow{۱\times۱۰=۱۰}$ صدک دهم $(P_{۱۰})$
- دهک پنجم : $D_۵ \xrightarrow{۵\times۱۰=۵۰}$ صدک پنجاهم $(P_{۵۰})$
- دهک ششم : $D_۶ \xrightarrow{۶\times۱۰=۶۰}$ صدک شصتم $(P_{۶۰})$



ایستگاه خنده:

حکایات شیرین بهلول:

یکـی از مســتخدمین خلیفـه هـارون الرشــید، ماســت خورده بود و قدری از آن بر ریشش چسبیده بود. بهلـول از او سوال کرد : چه خورده ای؟

مستخدم با تمسخر گفت : کبوترخورده ام.

بهلول گفت : قبل از آن که بگویی من میدانستم.

مستخدم پرسید : از کجا میدانستی؟

بهلول گفت : چون فضله آن بر ریشت پیدا بود.

«نحوهٔ محاسبهٔ چندک‌ها در داده‌های نوع اول (حجم کم)»

مثال: چارک سوم، دهک ششم و صدک هفتاد و هشتم را در مشاهدات زیر محاسبه کنید. (مفهوم این چندک‌های محاسبه شده چیست؟)

۲۰ و ۱۰۰ و ۹۰ و ۷۳ و ۴۵ و ۸۴ و ۳۰ و ۱۵ و ۱۰ و ۷۳

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۳۷ و ۳۸

«محاسبه چندک‌ها در داده‌های خام (نوع اول: حجم کم)»

گام ۱) *مرتب کردن داده‌ها به صورت صعودی و کدگذاری اونا از ۱ تا* N

گام۲) یافتن **محل چندک** (چارک، دهک و صدک) مورد نظر از رابطه زیر:

۳ و =a : شماره چارک ؛ $C_{Q_a}=\frac{aN}{۴}+\frac{۱}{۲}$ = **محل چارک a ام**

۹و ... و ۲و ۱ =a : شماره دهک ؛ $C_{D_a}=\frac{aN}{۱۰}+\frac{۱}{۲}$ = **محل دهک a ام**

۹۹و ... و ۲و ۱ =a : شماره صدک ؛ $C_{P_a}=\frac{aN}{۱۰۰}+\frac{۱}{۲}$ = **محل صدک a ام**

گام۳) یافتن **مقدار چندک:** داده متناظر با محل چندک را از بین داده-های مرتب شده بدست می‌آوریم.

مثال: *چارک سوم، دهک ششم و صدک هفتاد و هشتم را در مشاهدات زیر محاسبه کنید:* *۲۰ و ۱۰۰ و ۹۰ و ۷۳ و ۸۵ و ۸۴ و ۳۰ و ۱۵ و ۱۰ و ۷۳*

گام۱) صعودی کردن داده‌ها و کدگذاری اون:

(۱۰) (۹) (۸) (۷) (۶) (۵) (۴) (۳) (۲) (۱) : کد

۱۰۰ و ۹۰ و ۸۴ و ۷۳ و ۷۳ و ۴۵ و ۳۰ و ۲۰ و ۱۵ و ۱۰: مشاهدات

(ادامه تو فیش بعد)

گام۲) یافتن محل چندک مورد نظر:

$$C_{Q_۳} = \frac{۳N}{۴} + \frac{۱}{۲} \xrightarrow{N=۱۰} \frac{۳\times ۱۰}{۴} + \frac{۱}{۲} \Rightarrow$$ **محل چارک سوم**

$$\frac{۳۰}{۴} + \frac{۱}{۲} \Rightarrow \frac{۱۵}{۲} + \frac{۱}{۲} \Rightarrow ۷/۵ + ۰/۵ = ۸ :$$ ۸اُمین داده

خب، با توجه به کدهایی که بالای داده‌هامون نوشته‌ایم، کاملاً معلومه که ۸اُمین داده، $x = ۸۴$ است، پس:

$Q_۳ = ۸۴$: **چارک سوم**

$$C_{D_۶} = \frac{۶N}{۱۰} + \frac{۱}{۲} \xrightarrow{N=۱۰} \frac{۶\times \cancel{۱۰}}{\cancel{۱۰}} + \frac{۱}{۲} \Rightarrow$$ **محل دهک ششم**

$$۶ + ۰/۵ \Rightarrow$$ محل دهک ششم : $۶/۵ \Rightarrow$ ٦اُمین داده

با نگاه به کد مشاهدات‌مون می‌بینیم که ۶اُمین داده، $x = ۷۳$ است، پس:

$D_۶ = ۷۳$: **دهک ششم**

$$C_{P_{۷۸}} = \frac{۷۸N}{۱۰۰} + \frac{۱}{۲} = \frac{۷۸\times \cancel{۱۰}}{\cancel{۱۰۰}} + \frac{۱}{۲} \Rightarrow$$ **محل صدک هفتاد و هشتم**

$$\Rightarrow \frac{۷۸}{۱۰} + \frac{۱}{۲} \Rightarrow ۷/۸ + ۰/۵ = ۸/۳ \Rightarrow$$ ۸/۳اُمین داده

نکته مهم: برای یافتن مقدار ۸/۳اُمین داده باید اول دادهٔ ۸اُم (یعنی ۸۴) رو در نظر بگیریم و بعد به اندازهٔ ۰/۳ فاصله بین دادهٔ هشتم یا نهم، به سمت راست حرکت کنیم،

(ادامه تو پشت فیش)

یعنی باید ۰/۳ فاصله $x_{۸} = ۸۴$ و $x_{۹} = ۹۰$ رو به دادهٔ ۸اُم که ۸۴ است اضافه کنیم تا به دادهٔ ۸/۳ اُم برسیم، یعنی بصورت زیر:

$$\text{۸/۳ اُمین داده} = \text{۸اُم دادهٔ} + ۰/۳\,(\text{فاصله دادهٔ ۸اُم تا ۹اُم})$$
$$x_{۸/۳} = x_{۸} + ۰/۳(x_{۹} - x_{۸})$$

$$\text{۸/۳اُمین داده}: x_{۸/۳} = ۸۴ + ۰/۳(۹۰ - ۸۴) \Rightarrow$$
$$= ۸۴ + ۰/۳(۶) \Rightarrow$$
$$۸۴ + ۱/۸ = ۸۵/۸$$

پس:

$$P_{۷۸} = ۸۵/۸ : \textbf{صدک هفتاد و هشتم}$$

(GIS ۸۰)

در داده‌های آماری ۲۴ و ۵۹ و ۱۱ و ۴۱ و ۷ و ۳۵ چارک دوم کدام است؟

۱) ۱۵

۲) ۲۴

۳) ۴۱

۴) ۵۹

پاسخ: گزینه ۲

چارک دوم همان میانه است: $Q_۲ = Md$

پس به جای محاسبه چارک دوم، می‌تونیم میانه رو حساب کنیم چون حساب کردنش راحت‌تر و سریع‌تره:

گام ۱) صعودی کردن داده‌ها و کدگذاری اون:

کد : (۱) (۲) (۳) (۴) (۵) (۶) (۷) (۸)

x_i : ۱۴ و ۱۵ و ۱۶ و $\underbrace{\text{۱۷ و ۱۸}}_{\text{دو داده وسط}}$ و ۲۶ و ۲۷ و ۳۰

گام۲) تعیین محل میانه: چون تعداد داده‌ها زوجه $(N = ۸)$، پس حتماً ۲ داده، در وسط توزیع قرار می‌گیرن که میانه، برابر با **میانگین این دو دادهٔ وسط** خواهد بود.

گام ۳) تعیین مقدار میانه:

$$\text{میانه} = \text{میانگین ۲ دادهٔ وسط} \Rightarrow Md = \frac{x_۴ + x_۵}{۲} = \frac{۱۷+۱۸}{۲} = ۱۷/۵$$

«محاسبه چندک‌ها در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با داشتن فراوانی مطلق (F_i)»

مثال : *چارک اول در جدول داده‌های زیر کدام است؟*

(تفسیر آن را بیان کنید).

x_i	۴	۱	۰	۷
F_i	۹	۳	۴	۱۰

۱) ۱ ۲) ۴

۳) ۰ ۴) ۰/۲۵

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۸ و ۳۹

«**محاسبه چندک‌ها در داده‌های نوع دوم** (با داشتن فراوانی مطلق F_i)»

گام ۱) صعودی کردن طبقات و محاسبه **فراوانی تجمعی** طبقات (F_{C_i}).

گام۲) یافتن محل چندک از رابطه زیر:

$$\begin{cases} \textbf{محل چارک a ام} = C_{Q_a} = \frac{aN}{۴} + \frac{۱}{۲} \\ \textbf{محل دهک a ام} = C_{D_a} = \frac{aN}{۱۰} + \frac{۱}{۲} \\ \textbf{محل صدک a ام} = C_{P_a} = \frac{aN}{۱۰۰} + \frac{۱}{۲} \end{cases}$$

گام۳) یافتن مقدار چندک: از چپ به راست اولین طبقه (x_i)ای را انتخاب می‌کنیم که:

$$\begin{cases} ۱) \text{ برای چارکها ،} \quad F_{C_i} \geq \frac{aN}{۴} + \frac{۱}{۲} \text{ باشد.} \\ ۲) \text{ برای دهکها ،} \quad F_{C_i} \geq \frac{aN}{۱۰} + \frac{۱}{۲} \text{ باشد.} \\ ۳) \text{ برای صدکها ،} \quad F_{C_i} \geq \frac{aN}{۱۰۰} + \frac{۱}{۲} \text{ باشد.} \end{cases}$$

حال برای درک بهتر مراحل فوق، با ذکر ۳ مثال، نحوه محاسبه انواع چندک (چارک، دهک و صدک) را نشان می‌دهیم:

(ادامه در فیش بعد)

مثال۱: چارک اول در جدول داده‌های زیر کدام است؟

x_i	۴	۱	۰	۷	
F_i	۱	۸	۷	۲	$N = ۲۶$

پاسخ: گزینه ۱

گام۱) صعودی کردن طبقات **X**ها و محاسبه فراوانی تجمعی طبقات:

گام۲) یافتن محل چارک اول:

$$\frac{۱N}{۴} + \frac{۱}{۲} \Rightarrow \xrightarrow{\text{چارک اول } (a=۱)} \boxed{C_{Q_۱} = \frac{aN}{۴} + \frac{۱}{۲} : \text{محل چارک a ام}}$$

$$\frac{۱ \times ۲۶}{۴} + \frac{۱}{۲} = \frac{۱۳}{۲} + \frac{۱}{۲} = ۶ / ۵ + ۰ / ۵ = ۷ \rightarrow \boxed{\text{۷ اُمین داده}}$$

(داده ۵ ام تا ۷ ام) (داده یکم تا ۴ اُم)

x_i	۰	۱	۴	۷	
F_i	۴	۳	۹	۱۰	$\Sigma F_i = N = ۲۶$
F_{C_i}	۴	۷	۱۶	۲۶	

گام۳) یافتن مقدار چندک: اولین طبقه‌ای که در آن فراوانی تجمعی بزرگتر یا مساوی محل چارک اول باشه، یعنی اولین طبقه‌ای که فراوانی تجمعیش **بزرگتر یا مساوی** ۷ باشه $(F_{C_i} \geq ۷)$، طبقه $x = ۱$ است، پس:

$$\boxed{\text{چارک اول} = \text{داده ۷اُم} \Rightarrow Q_۱ = ۱}$$

روش دوم:

می تونیم به این صورت هم بگیم که با توجه به جدول، از دادهٔ ۵اُم تا ۷اُم تو طبقه $x=۱$ قرار دارن، پس چارک اول هم که دادهٔ ۷ اُم ماست، تو همین طبقه $(x=۱)$ قرار داره: $\boxed{Q_۱ = ۱}$

مفهومِ $Q_۱ = ۱$: اینه که $\frac{۱}{۴} = ۲۵٪$ داده‌ها **کمتر یا مساوی** $x = ۱$ هستن و ۷۵٪ یا $\frac{۳}{۴}$ داده‌ها **بزرگتر از** $x = ۱$ هستن.

«محاسبه چندک‌ها در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با داشتن فراوانی مطلق (F_i)»

مثال ۲: دهک ششم در جدول داده‌های زیر کدام است؟

(آن را تفسیر کنید)

x_i	۴	۱	۰	۷
F_i	۷	۵	۶	۸

۱) ۰/۶ ۲) ۴

۳) ۲/۴ ۴) ۰

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۸ و ۳۹

پاسخ: گزینه ۲

گام۱) صعودی کردن طبقات و محاسبه فراوانی تجمعی:

گام۲) یافتن محل دهک ششم:

$$\frac{۶N}{۱۰}+\frac{۱}{۲}\Rightarrow \xrightarrow{\text{دهک ۶ اُم } (a=۶)} \boxed{\text{محل دهک ششم}=C_{Q_۶}=\frac{aN}{۴}+\frac{۱}{۲}}$$

$$\xrightarrow{N=۲۶}\frac{۶\times ۲۶}{۱۰}+\frac{۱}{۲}\Rightarrow\frac{۱۵۶}{۱۰}+\frac{۱}{۲}=۱۵/۶+۰/۵=\boxed{۱۶/۱\text{ اُمی}}$$

(داده ۱ اُم تا ۶ اُم) (داده ۷ اُم تا ۱۱ اُم) (داده ۱۲ اُم تا ۱۸ اُم)

x_i	۰	۱	۴	۷	
F_i	۶	۵	۷	۸	$\sum F_i = N = ۲۶$
F_{C_i}	۶	۱۱	۱۸	۲۶	

گام۳) یافتن مقدار دهک ششم: اولین طبقه‌ای که فراوانی تجمعیش بزرگتر یا مساوی محل دهک ششم است (یعنی $F_{C_i} \geq ۱۶/۱$)، طبقه $x=۴$ است:

$$\xrightarrow{\text{در طبقه } x=۴} F_{C_۳}=۱۸\geq ۱۶/۱\Rightarrow \boxed{\text{دهک ششم}: \text{داده ۱۶/۱ اُم} \Rightarrow D_۶=۴}$$

روش دوم: با توجه به جدول می‌بینیم که از دادهٔ ۱۲اُم تا ۱۸اُم تو طبقه $x=۴$ قرار دارن، پس مسلماً دهک ششم که دادهٔ ۱۶/۱اُم ماست هم، تو همین طبقه $(x=۴)$ قرار داره: $\boxed{D_۶=۴}$

مفهومِ $D_۶=۴$: این است که $\frac{۴}{۱۰}=۴۰\%$ داده‌ها **کوچکتر یا مساوی** $x=۴$ و $\frac{۶}{۱۰}=۶۰\%$ داده‌ها **بزرگتر از** $x=۴$ هستند.

«محاسبه چندک‌ها در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با داشتن فراوانی مطلق (F_i)»

مثال ۳: *صدک ۳۶ اُم در جدول داده‌های زیر کدام است؟*

(مفهوم این عدد چیست؟)

x_i	۴	۱	۰	۷
F_i	۱۶	۵	۹	۱۰

۱) ۰

۲) ۰/۳۶

۳) ۲/۵

۴) ۳/۷

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۳۸ و ۳۹

پاسخ: گزینه ۴

گام۱) **صعودی** کردن طبقات و محاسبه فراوانی **تجمعی** طبقات (F_{C_i})

گام۲) یافتن **محل** صدک سی و ششم:

$$\boxed{\text{محل صدک ۳۶ اُم} = C_{Q_{۳۶}} = \frac{aN}{۱۰۰} + \frac{۱}{۲}} \xrightarrow{a=۳۶} \frac{۳۶N}{۱۰۰} + \frac{۱}{۲} \Rightarrow$$

$$\xrightarrow{N=۴۰} \frac{۳۶ \times ۴۰}{۱۰۰} + \frac{۱}{۲} \Rightarrow \frac{۱۴۴۰}{۱۰۰} + \frac{۱}{۲} \Rightarrow ۱۴/۴ + ۰/۵ = \boxed{\text{۱۴/۹ اُمی}}$$

(داده ۵ اُم تا ۳۰ اُم) (داده ۰ اُم تا ۴ اُم) (داده ۱ اُم تا ۹ اُم)

x_i	۰	۱	۴	۷	
F_i	۹	۵	۱۶	۱۰	$\sum F_i = N = ۴۰$
F_{C_i}	۹	۱۴	۳۰	۴۰	

گام۳) یافتن **مقدار** صدک سی و ششم:

اولین طبقه (x_i)ای که فراوانی تجمعی‌اش بزرگتر یا مساوی محل صدک سی و ششم است (یعنی $F_{C_i} \geq ۱۴/۹$)، طبقه $x=۴$ است، اما باید دقت کنیم که در این جا نمی‌تونیم بگیم که صدک سی و ششم در طبقه $x=۴$ قرار داره، **(ادامه تو فیش بعد)**

یعنی نمی‌تونیم بگیم که صدک سی و ششم مساوی $x = ۴$ است، چون با نگاه به جدول می‌فهمیم که از دادهٔ ۱۰ اُم تا ۱۴ اُم تو طبقه $x = ۱$ قرار داره و از دادهٔ ۱۵ اُم تا ۳۰ اُم هم تو طبقه $x = ۴$ قرار داره، بنابراین دادهٔ ۱۴/۹ اُم، بیچاره مثل یه بچه بی سرپرسته که نه تو طبقه $x = ۱$ (۱۰ اُم تا ۱۴ اُم) جاشه و نه تو طبقه $x = ۴$ (۱۵ اُم تا ۳۰ اُم).

یعنی دادهٔ ۱۴/۹ اُم تو برزخ قرار گرفته، یعنی بین دادهٔ ۱۴ اُم (که مربوط به طبقه $x = ۱$ است) و دادهٔ ۱۵ اُم (که مربوط به طبقهٔ $x = ۴$ است)، قرار گرفته، بنابراین دادهٔ ۱۴/۹ اُم رو باید بصورت زیر حساب کنیم:

(فاصله دادهٔ ۱٤ اُم تا ۱۵ اُم) ۰/۹ + دادهٔ ۱٤ اُم = دادهٔ ۱٤/۹ اُم

$$\Rightarrow x_{۱۴/۹} = x_{۱۴} + ۰/۹\,(x_{۱۵} - x_{۱۴})$$

$$= ۱ + ۰/۹(۴ - ۱) \Rightarrow$$

$$= ۱ + ۰/۹(۳) \Rightarrow ۱ + ۲/۷ = ۳/۸$$

یعنی: دادهٔ ۱۴/۹ اُم $\Rightarrow P_{۳۶} = ۳/۸$ = **صدک ۳۶ اُم**

مفهوم $P_{۳۶} = ۳/۷$: این است که ۳۶٪ = $\frac{۳۶}{۱۰۰}$ مشاهدات **کوچکتر یا مساوی** $x = ۳/۷$ هستن و ۶۴٪ = $\frac{۶۴}{۱۰۰}$ مشاهدات باقیمانده، **بزرگتر از** $x = ۳/۷$ هستند.

آنکس که بداند و بداند که بداند — اسب خرد از گنبد گردون بجهاند
آنکس که بداند و نداند که بداند — بیدارش نمایید که بس خفته نماند
آنکس که نداند و بداند که نداند — لنگان خرک خویش به منزل برساند
آنکس که نداند و نداند که نداند — در جهل مرکب ابدالدهر بماند
آنکس که نداند و نخواهد که بداند — حیف است چنین جانوری زنده بماند

«نحوه محاسبه چندک‌ها در داده‌های نوع دوم (حجم زیاد، تنوع کم)

با داشتن فراوانی نسبی: (f_i)»

مثال: چارک سوم در جدول داده‌های زیر کدام است؟

(این عدد به چه مفهومی است؟)

x_i	۴	۱	۰	۷
f_i	۰/۳	۰/۱	۰/۴	۰/۲

۱) ۰ ۲) ۴

۳) ۲/۵ ۴) ۱

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ٤٠

«**محاسبه چندک‌ها در داده‌های نوع دوم (با داشتن فراوانی نسبی: f_i)**»

گام ۱) **صعودی** کردن طبقات و محاسبهٔ فراوانی تجمعی **نسبی** طبقات (f_i)

گام ۲) **یافتن محل چندک:**

اولین طبقه (x_i)ای که فراوانی تجمعی نسبی آن به ترتیب:

۱) برای **چارک aام** بزرگتر یا مساوی $\frac{a}{۴}$ باشد ($f_{C_i} \geq \frac{a}{۴}$)
مثال: برای چارک سوم باید $f_{C_i} \geq \frac{۳}{۴} = ۰/۷۵$ باشد.
۲) برای **دهک aام** بزرگتر یا مساوی $\frac{a}{۱۰}$ باشد ($f_{C_i} \geq \frac{a}{۱۰}$)
مثال: برای دهک ششم ($a = ۶$)، باید: $f_{C_i} \geq \frac{۶}{۱۰} = ۰/۶$
۳) برای **صدک aام** بزرگتر یا مساوی $\frac{a}{۱۰۰}$ باشد ($f_{C_i} \geq \frac{a}{۱۰۰}$)
مثال: برای صدک بیستم ($a = ۲۰$)، باید: $f_{C_i} \geq \frac{۲۰}{۱۰۰} = ۰/۲۰$

(ادامه تو فیش بعد)

پاسخ: گزینه ۲

گام ۱) صعودی کردن طبقات و محاسبه فراوانی تجمعی نسبی

(در جدول زیر):

x_i	۰	۱	۴	۷	
f_i	۰/۴	۰/۱	۰/۳	۰/۲	$\sum f_i = ۱$
f_{C_i}	۰/۴	۰/۵	۰/۸	۱	

گام ۲) یافتن محل چارک سوم:

اولین طبقه‌ای که فراوانی تجمعیش بزرگتر از $\frac{a}{۴} = \frac{۳}{۴} = ۰/۷۵$ باشه، طبقه $x = ۴$ است، پس چارک سوم ما مساوی $x = ۴$ است:

$$\xrightarrow{x=۴} f_C = ۰/۸ \geq ۰/۷۵ \Rightarrow \boxed{Q_۳ = ۴ : \text{چارک سوم}}$$

توجه کنین که: *از اونجایی که چارک سوم دقیقاً منطبق با صدک هفتاد و پنجمه، پس اگه تو این سؤال از ما مقدار $P_{۷۵}$ رو پرسیده بودن، باز هم جواب ما همون $x = ۴$ می‌شد:*

یعنی: $\boxed{Q_۳ = P_{۷۵} = ۴}$

(ادامه تو پشت فیش)

مفهوم $Q_{۳} = ۴$: اینه که ٪۷۵ $= \frac{۳}{۴}$ داده‌ها **کوچکتر یا مساوی** $x = ۴$ هستن و ٪۲۵ $= \frac{۱}{۴}$ مشاهدات باقیمانده، **بزرگتر از** $x = ۴$ هستند.

(بسیار مهم):

«نحوهٔ محاسبه چندک‌ها در داده‌های نوع سوم

(حجم زیاد, تنوع زیاد) با داشتن فراوانی مطلق : (F_i)»

مثال) (طبقات پیوسته): *چارک اول داده‌های زیر کدام است؟*

(مفهوم آن چیست؟)

C-L	۰-۶	۶-۱۲	۱۲-۱۸	۱۸-۲۴
F_i	۳	۱۰	۷	۴

۱) ۷/۸ ۲) ۱۲

۳) ۱/۸ ۴) ۶/۳

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ٤١

«**محاسبه چندک‌ها در داده‌های نوع سوم (دارای طبقه‌بندی فاصله‌ای)**»

توجه کنین که: *در این حالت که طبقه‌بندی ما به صورت فاصله‌ای است، یعنی طبقات به شکل (حد بالا – حد پایین) هستن، برای محاسبهٔ چندک‌ها دیگه لازم نیست طبقات رو به صورت* ***صعودی*** *مرتب کنیم، چون قاعدهٔ طبقه‌بندی کردن به صورت فاصله‌ای به این صورته که طبقات حتماً باید به صورت صعودی مرتب بشن.*

بنابراین چون در این حالت، طبقات ما از قبل، به صورت صعودی مرتب شده‌اند،پس دیگه نیازی نیست که ما دوباره این کار رو انجام بدیم.

گام ۱) محاسبه فراوانی **تجمعی** (F_{C_i}) طبقات:

گام ۲) یافتن **محل** چندک (طبقه چندک دار):

اولین طبقه (از چپ به راست) که فراوانی تجمعیش بیشتر یا مساوی $\frac{aN}{۴}$ یا $\frac{aN}{۱۰}$ یا $\frac{aN}{۱۰۰}$ باشد، یعنی:

($F_{C_i} \geq \frac{aN}{۴}$ و $\frac{aN}{۱۰}$ و $\frac{aN}{۱۰۰}$)، بترتیب نشون دهندهٔ طبقه چارک دار، دهک دار یا صدک دار ما است. **(ادامه تو فیش بعد)**

گام۳) پیوسته کردن طبقه چندک دار در صورت گسسته بودن.

گام۴) یافتن مقدار تقریبی چندک از روابط زیر:

$$\text{چارک a ام}: Q_a \approx L_i + \frac{\frac{aN}{۴} - F_{C_{i-۱}}}{F_i} \times I \quad ؛ \quad a = ۱,۲,۳$$

$$\text{دهک a ام}: D_a \approx L_i + \frac{\frac{aN}{۱۰} - F_{C_{i-۱}}}{F_i} \times I \quad ؛ \quad a = ۱,۲,...,۹$$

$$\text{صدک a ام}: P_a \approx L_i + \frac{\frac{aN}{۱۰۰} - F_{C_{i-۱}}}{F_i} \times I \quad ؛ \quad a = ۱,۲,...,۹۹$$

که در روابط فوق:

L_i : کران پایین **واقعی** طبقه چندک‌دار

$F_{C_{i-۱}}$:فراوانی تجمعی طبقه **قبل از** طبقه چندک‌دار

F_i :فراوانی مطلق طبقه چندک‌دار

I :طول طبقات = فاصله طبقات = عرض طبقات

(ادامه تو پشت فیش)

توجه: چون تو گام۳، اول اومدیم طبقات رو پیوسته کردیم، بنابراین با توجه به **پیوسته** بودن طبقات می‌تونیم بگیم که طول، عرض و فاصله طبقات با هم برابرند).

مثال((طبقات پیوسته): *چارک اول داده‌های زیر کدام است؟*

پاسخ: گزینه یک

گام ۱) محاسبه فراوانی تجمعی (در جدول زیر):

(داده ۴(اُم تا ۳(اُم) (داده یکم تا ۳(اُم)

$C-L$	۰-۶	۶-۱۲	۱۲-۱۸	۱۸-۲۴	
F_i	۳	۱۰	۷	۴	$\Sigma F_i = N = ۲۴$
F_{C_i}	۳	۱۳	۲۰	۲۴	

گام۲) یافتن محل چارک اول: اولین طبقه (از چپ به راست) که فراوانی تجمعیش از $\frac{aN}{۴}$ بزرگتره، طبقه دومه، پس این طبقه (۱۲-۶) طبقه چارک دار ماست:

(ادامه تو فیش بعد)

محل چارک اول : $\frac{aN}{۴} \xrightarrow{\text{چارک اول}(a=1)}$

$\frac{۱\times N}{۴} \xrightarrow{N=۲۴} \frac{۲۴}{۴} =$ **دادهٔ ۶اُم**

تعیین طبقه چارک‌دار (روش اول):

$\xrightarrow{\text{طبقه دوم}} F_{C_۲} = ۱۳ \geq ۶ \overset{\text{پس}}{\Longrightarrow}$ طبقه دوم (۱۲-۶) = **طبقه چندک دار**

تعیین طبقه چارک‌دار (روش دوم):

با توجه به جدول می‌بینیم که از دادهٔ اول تا سوم در طبقه (۶-۰) قرار گرفته‌اند و از دادهٔ ۴ اُم تا ۱۳ اُم هم تو طبقه (۱۲-۶) جای گرفته‌اند؛ پس می‌تونیم بگیم که دادهٔ ۶ اُم (که همون چارک اول ماست) هم حتماً تو همین طبقه (۱۲-۶) قرار داره.

توجه کنین که: تو این مثال، چون طبقات ما **پیوسته** بودن، پس دیگه نیازی به انجام گام ۳ (یعنی **پیوسته** کردن طبقه چندک‌دار) نبود.

گام۳) یافتن **مقدار** تقریبی چارک اول از رابطه زیر:

$$\boxed{\textbf{مقدار چارک}: Q_۱ \approx L_i + \frac{\frac{۱\times N}{۴} - F_{C_{i-۱}}}{F_i} \times I} \Rightarrow ۶ + \frac{\frac{۲۴}{۴} - ۳}{۱۰} \times ۶ \Rightarrow$$

(ادامه تو پشت فیش)

$$۶+\frac{۶-۳}{۱۰}\times ۶ \Rightarrow ۶+\frac{۳\times ۶}{۱۰} \Rightarrow ۶+\frac{۱۸}{۱۰} \Rightarrow ۶+۱/۸ \Rightarrow$$

$\boxed{\text{چارک اول}: Q_۱ = ۷/۸}$

مفهومِ $Q_۱ = ۷/۸$: این است که ۲۵٪ $= \frac{۱}{۴}$ مشاهدات **کوچکتر یا مساوی** $x = ۷/۸$ هستن و ۷۵٪ $= \frac{۳}{۴}$ مشاهدات **بزرگتر از** $x = ۷/۸$ هستن.

(بسیار مهم):

«نحوهٔ محاسبه چندک‌ها در داده‌های نوع سوم (حجم زیاد، تنوع زیاد)

با داشتن فراوانی مطلق : (F_i) »

مثال۲ (طبقات گسسته): *در جدول زیر صدک ۶۸اُم کدام است؟*

(مفهوم آن را بیان کنید).

C-L	۴-۶	۷-۹	۱۰-۱۲	۱۳-۱۵
F_i	۸	۷	۱۰	۱۲

۱) ۱۳/۱۹ ۲) ۱۲/۵۴

۳) ۱۲/۱۲ ۴) ۱۰/۷۵

آمار و احتمالات، آموزشگاه مدرسان شریف، ص ۲۰ مثال ۷۵

پاسخ: گزینه ۲

گام۱) محاسبه فراوانی تجمعی طبقات (در جدول).

گام۲) یافتن محل صدک ۶۸اُم: اولین طبقه‌ای که فراوانی تجمعیش **بزرگتر یا مساوی** $\frac{aN}{۱۰۰}=\frac{۶۸N}{۱۰۰}$ باشد ($F_{C_i} \geq \frac{۶۸N}{۱۰۰}$)، طبقه چهارمه (۱۵-۱۳):

محل صدک ۶۸اُم: $\frac{aN}{۱۰۰} \xrightarrow{a=۶۸} \frac{۶۸N}{۱۰۰} \xrightarrow{N=۳۷} \frac{۶۸\times ۳۷}{۱۰۰}=\frac{۲۵۱۶}{۱۰۰}$

محل صدک ۶۸اُم = ۲۵ / ۱۶ $\Rightarrow$ دارهٔ ۲۵/۱۶ اُمی

تعیین طبقه چندک‌دار :

طبقه (۱۵-۱۳) = طبقه چندک‌دار $\xrightarrow{\text{پس}}$ $F_C = ۳۷ \geq ۲۵ / ۱۶ \xrightarrow{\text{در طبقه چهارم}}$

اما یه نکته مهم: واسهٔ حساب کردن صدک ۶۸اُم خیلی عجله نکنین، چون همون‌طور که تو جدول می‌بینین، طبقات ما به صورت **گسسته** هستن، پس ما **اول باید طبقات رو پیوسته کنیم** که البته برای سرعت عمل بیشتر، فقط کافیه که طبقه چندک‌دار، یعنی طبقه (۱۵-۱۳) رو **پیوسته** کنیم:

(ادامه تو فیش بعد)

گام۳) پیوسته کردن طبقه صدک‌دار:

$$(۱۳-۱۵)\xrightarrow{\text{حدود واقعی}}(۱۳-۰/۵\ ,\ ۱۵+۰/۵)=(۱۲/۵-۱۵/۵)$$

C-L	۴-۶	۷-۹	۱۰-۱۲	۱۳-۱۵	
حدود واقعی	۳/۵-۶/۵	۶/۵-۹/۵	۹/۵-۱۲/۵	۱۲/۵-۱۵/۵	
F_i	۸	۷	۱۰	۱۲	$\Sigma F_i = N = ۳۷$
F_{C_i}	۸	(۷+۸) ۱۵	(۱۰+۱۵) ۲۵	(۱۲+۲۵) ۳۷	

گام۴) یافتن **مقدار** تقریبی صدک ۶۸اُم از رابطه زیر:

$$\text{مقدار صدک ۶۸اُم} = P_{۶۸} \cong L_i + \frac{\frac{۶۸\times N}{۱۰۰} - F_{C_{i-۱}}}{F_i}\times I \Rightarrow$$

$$P_{۶۸} = ۱۲/۵ + (\frac{۲۵/۱۶-۲۵}{۱۲})\times ۳ \Rightarrow ۱۲/۵ + (\frac{۰/۱۶}{۱۲})\times ۳ \Rightarrow$$

$$P_{۶۸} = ۱۲/۵ + \frac{۰/۴۸}{۱۲} \Rightarrow ۱۲/۵ + ۰/۰۴ \Rightarrow \boxed{P_{۶۸} = ۱۲/۵۴}$$

(ادامه تو پشت فیش)

مفهومِ $P_{۶۸} = ۱۲/۵۴$: اینه کـه ۶۸٪ $= \frac{۶۸}{۱۰۰}$ مشـاهدات، **کوچکتر یا مساوی** $x = ۱۲/۵۴$ هستن و ۳۲٪ $= \frac{۳۲}{۱۰۰}$ مشاهدات هم **بزرگتـر از** $x = ۱۲/۵۴$ هستن.

نکته۱: *با توجه به ۲ مثال حل شده در بالا ملاحظه می‌شود که پیوسته یا گسسته بودن طبقات، تفاوتی در نحوهٔ محاسبه چندک‌ها ایجاد نمی‌کنه و تنها کار اضافی‌ای که موقع گسسته بودن طبقات باید انجام بدیم، اینه که قبل از محاسبه چندک، اول بیائیم* **طبقه چندک‌دار** *رو پیوسته کنیم.*

نکته۲: *تمام محاسبات مربوط به پارامترهای مرکزی (مد، میانه، میانگین و چندک‌ها) با فرض پیوسته بودن حدود طبقات انجام می‌شوند. یعنی* **در محاسبه تمام پارامترهای مرکزی باید طول (فاصله) و عرض طبقه مساوی باشد.** *بنابراین اگر حدود طبقات با تقریب (۱، ۰/۱، ۰/۰۱و...) صورت گرفته باشد، موقع حساب کردن مد، میانگین یا چندک مورد نظر، باید حتماً و حتماً از حدود* **واقعی** *طبقات (یعنی همون حدود* **کرانه‌ها**) *استفاده کنیم، یعنی اول باید خیالمون از پیوسته بودن طبقات راحت بشه و بعدش بریم سراغ حساب کردن شاخص مرکزی یا مکانی مورد نظر (مثل همون کاری که ما تو این ۲ مثال قبلی انجام دادیم، یعنی محاسباتمون رو روی طبقات* **پیوسته** *انجام دادیم).*

(محیط زیست ۸۴)

با توجه به جدول طبقه‌بندی، مقدار چارک سوم کدام است؟

C-L	۱۰-۲۰	۲۰-۳۰	۳۰-۴۰
F_i	۱۰	۳۰	۲۰

۱) ۲۵

۲) ۳۰/۵

۳) ۳۲/۵

۴) ۴۵

پاسخ: گزینه ۳

گام ۱) محاسبه فراوانی تجمعی طبقات (در جدول زیر):

C-L	داده اول تا دهم $\overbrace{10-20}$	داده ۱۰اُم تا ۴۰اُم $\overbrace{20-30}$	داده ۴۰اُم تا ۶۰اُم $\overbrace{\boxed{30}-40}$	
F_i	10	30	$\boxed{20}$	$\Sigma F_i = N = ۶۰$
F_{C_i}	10	$\boxed{40}$	60	

گام ۲) یافتن محل چارک سوم:

اولین طبقه‌ای که فراوانی تجمعیش بزرگتر یا مساوی $\frac{۳N}{۴} = ۴۵$ باشه $(F_{C_i} \geq ۴۵)$، طبقه سومه (۴۰-۳۰)، پس:

محل چارک سوم: $\frac{۳N}{۴} \xrightarrow{N=۶۰} \frac{۳ \times \cancel{۶۰}}{\cancel{۴}} = ۳ \times ۵ = ۴۵ \rightarrow$ $\boxed{\text{داده ۴۵اُم}}$

تعیین طبقه چارک‌دار (روش اول):

$\xrightarrow{\text{در طبقه سوم}} F_C = ۶۰ \geq ۴۵ \xrightarrow{\text{پس}}$

$\boxed{\text{طبقهٔ (۴۰-۳۰) : } \textbf{طبقه چارک‌دار}}$

تعیین طبقه چارک‌دار (روش دوم): با توجه به جدول بالا می‌بینیم که داده‌های ۴۰اُم تا ۶۰اُم تو طبقه (۴۰-۳۰) قرار دارن، پس نتیجه می‌گیریم

(ادامه تو فیش بعد)

که دادهٔ ۴۵اُم (که همون چارک سومه) هم تو همین طبقه (۴۰-۳۰) قرار داره.

توجه: تو این سؤال، طبقات ما **پیوسته** هستن، پس دیگه خدا رو شکر نیازی به **پیوسته** کردن طبقه **چارک‌دار** نداریم و می‌تونیم یه راست بریم سراغ گام ۳، یعنی محاسبه مقدار چارک:

گام ۳: محاسبه چارک سوم:

$$\text{مقدار چارک سوم}: Q_۳ \approx \text{حد پایین دسته} + \frac{\frac{۳N}{۴} - F_{C_{i-۱}}}{F_i} \times \text{طول دسته}$$

$$Q_۳ \approx ۳۰ + \frac{۴۵ - ۴۰}{۲۰} \times ۱۰ \Rightarrow$$

$$Q_۳ \approx ۳۰ + \frac{۵}{\cancel{۲۰}} \times \cancel{۱۰} \Rightarrow$$

$$Q_۳ \approx ۳۰ + \frac{۵}{۲} \Rightarrow Q_۳ \approx ۳۰ + ۲/۵ \Rightarrow$$

$$\text{چارک سوم}: Q_۳ = ۳۲/۵$$

نکته تستی: *از همون لحظه‌ای که فهمیدیم چارک سوم تو طبقه (۴۰-۳۰) داره، می‌تونستیم گزینه‌های ۱ و ۴ رو حذف کنیم، چون مقدار اونا تو فاصله (۴۰-۳۰) قرار نداره.*

(مدیریت ۸۰)

در جدول زیر دهک دوم برابر چند است؟

C-L	۴۰-۵۰	۵۰-۶۰	۶۰-۷۰
F_i	۵	۱۸	۷

۱) ۴۸/۲

۲) ۵۱/۵

۳) ۵۰/۵۵

۴) ۶۲/۳۸

پاسخ: گزینه ۳

گام ۱) محاسبه فراوانی تجمعی طبقات (در جدول زیر):

C-L	40-50	٦ اُم تا ۲۳ اُم $\boxed{50} - 60$	60-70	
F_i	5	$\boxed{18}$	7	$\Sigma F_i = N = ۳۰$
F_{C_i}	$\boxed{5}$	23	30	

گام ۲) یافتن محل دهک دوم:

اولین طبقه‌ای که فراوانی تجمعیش **بزرگتر یا مساوی** $\frac{۲N}{۱۰} = ۶$ باشه $(F_{C_i} \geq ۶)$، طبقه دومه (۶۰-۵۰):

محل دهک دوم: $\frac{۲N}{۱۰} \xrightarrow{N=۳۰} \frac{۲\times ۳۰}{۱۰} = \frac{۶۰}{۱۰} = ۶ \Rightarrow$ دارهٔ ٦ اُم

یافتن طبقه دهک‌دار (روش اول): همون‌طور که تو جدول بالا می‌بینیم، داره‌های ۶ اُم تا ۲۳ اُم تو طبقه (۶۰-۵۰) قرار دارن، پس دارهٔ ۶ اُم (که همون دهک دومه) هم تو همین طبقه (۶۰-۵۰) قرار می‌گیره.

یافتن طبقه دهک‌دار (روش دوم):

$\xrightarrow{\text{در طبقه دوم}} F_C = ۲۳ \geq ۶ \xrightarrow{\text{پس}}$ (۶۰-۵۰) : **طبقه دهک‌دار**

(ادامه تو فیش بعد)

توجه: چون طبقات ما **پیوسته** هستن، پس دیگه لازم نیست طبقه دهک‌دار رو پیوسته کنیم.

$$\text{مقدار دهک دوم}: D_{۲} = \text{حد پایین طبقه} + \frac{\frac{۲N}{۱۰} - F_{C_{i-۱}}}{F_i} \times \text{فاصله (طول) طبقه}$$

$$D_{۲} = ۵۰ + \frac{۶-۵}{۱۸} \times ۱۰ \Rightarrow D_{۲} = ۵۰ + \frac{۱}{۱۸} \times ۱۰ \Rightarrow$$

$$D_{۲} = ۵۰ + \frac{\cancel{۱۰}}{\cancel{۱۸}} \Rightarrow D_{۲} = ۵۰ + \frac{۵}{۹} \Rightarrow$$

$$D_{۲} = ۵۰ + ۰/۵۵ \Rightarrow \boxed{D_{۲} = ۵۰/۵۵}$$

یه نکته ظریف: *تو گام (۲) فهمیدیم که دهک دوم تو طبقه (۶۰-۵۰) قرار داره، بنابراین با توجه به این مطلب، می‌تونستیم گزینه‌های ۱ و ۴ که مقدارشون در فاصله (۶۰-۵۰) نیست رو خیلی راحت حذف کنیم.*

واسهٔ حذف کردن یکی از دو گزینه باقیمونده (۲ یا ۳) هم روش‌های زیادی وجود داره که توسط بعضی از داوطلب‌ها انجام میشه. البته من این روش‌ها رو اصلاً به شما پیشنهاد نمی‌کنم. بعضی از این روش‌ها که شما بهتر از من باهاشون آشنایین اینا هستن:

- ده بیست سی چهل کردن
- تسبیح انداختن
- استخاره کردن و ...(بقیه شو خودتون بگین)؟؟!

(مدیریت ۷۸)

چارک اول جدول زیر کدام است؟

حدود طبقات	۲-۵	۶-۹	۱۰-۱۳
فراوانی مطلق	۱۰	۳۰	۲۰

۱) ۵

۲) ۶/۱۷

۳) ۹/۵

۴) ۱۰

پاسخ: گزینه ۲

گام ۱) محاسبه فراوانی تجمعی طبقات:

C-L	$2-5$	$\overbrace{6-9}^{\text{دادهٔ ۱۱ تا ۴۰ اُم}}$	$10-13$	
F_i	10	$\boxed{30}$	20	$\Sigma F_i = N = ۶۰$
F_{C_i}	$\boxed{10}$	40	60=N	

گام ۲) یافتن محل چارک اول:

اولین طبقه‌ای که فراوانی تجمعیش **بزرگتر یا مساوی** $\frac{۱N}{۴} = ۱۵$ باشد، $(F_{C_i} \geq ۱۵)$، طبقه (۹-۶) است:

محل چارک اول: $\frac{N}{۴} \xrightarrow{N=۶۰} \frac{۶۰}{۴} = ۱۵ \rightarrow$ دادهٔ ۱۵اُم

یافتن طبقه چارک‌دار (روش اول):

با توجه به جدول بالا می‌بینیم که داده‌های ۱۱ اُم تا ۴۰ اُم تو طبقه (۹-۶) قرار دارن، پس دادهٔ ۱۵ اُم (یعنی همون چارک اول) هم تو این طبقه (۹-۶) قرار می‌گیره. **(ادامه تو فیش بعد)**

یافتن طبقه چارک‌دار (روش دوم):

$$\xrightarrow{\text{در طبقه دوم}} F_C = ۴۰ \geq ۱۵ \xrightarrow{\text{پس}} \boxed{(۹-۶) : \textbf{طبقه چارک‌دار}}$$

گام۳) پیوسته کردن طبقه چارک‌دار:

$$(۹-۶) \xrightarrow{\pm ۰/۵} (۹+۰/۵ \text{ و } ۶-۰/۵) \Rightarrow \boxed{(۹/۵ \text{ و } ۵/۵)}$$

خب، حالا بعد از پیوسته کردن طبقه چارک‌دار می‌تونیم فاصله طبقات (I) رو بدست بیاریم:

$$\boxed{I = ۹/۵ - ۵/۵ = ۴}$$

گام ۴) محاسبه چارک اول:

$$\boxed{Q_۱ \approx \text{حد پایین } \textbf{واقعی} \text{ طبقه} + \frac{\frac{N}{۴} - F_{C_{i-۱}}}{F_i} \times \text{فاصله (طول) طبقه}}$$

$$Q_۱ \approx ۵/۵ + \frac{۱۵-۱۰}{۳۰} \times ۴ \Rightarrow Q_۱ \approx ۵/۵ + \frac{۵}{۳۰} \times ۴ \Rightarrow$$

$$Q_۱ \approx ۵/۵ + \frac{۲۰}{۳۰} \Rightarrow Q_۱ \approx ۵/۵ + \frac{۲}{۳} \Rightarrow$$

$$Q_۱ \approx ۵/۵۰ + ۰/۶۷$$

$$\boxed{\textbf{چارک اول} : Q_۱ \approx ۶/۱۷}$$

بسیار مهم:

(GIS ۸۷)

در ۱۲۰ داده‌ی آماری، کوچکترین و بزرگترین آنها به ترتیب ۱۲ و ۵۴ می‌باشد. این داده‌ها در ۷ طبقه دسته‌بندی شده‌اند، به طوری که مقدار دهک ششم برابر ۳۲ و در دسته وسط واقع است. اگر فراوانی مطلق این طبقه ۹ باشد، درصد فراوانی نسبی تجمعی آن کدام است؟

۱) ۶۳

۲) ۶۵

۳) ۶۸

۴) ۷۲

جواب: گزینه «۲»

گام۱) نوشتن داده‌های مسئله:

$$\begin{cases} N = ۱۲۰ : \text{تعداد داده‌ها} \\ X_{\max} = ۵۴ : \text{بزرگترین داده و } X_{\min} = ۱۲ : \text{کوچکترین داده} \\ K = ۷ : \text{تعداد طبقات} \\ D_۶ = ۳۲ \rightarrow : \text{دهک ششم} \quad \text{طبقه } \textbf{چهارم} = \text{طبقه } \textbf{وسط} = \text{طبقه دهک‌دار} \\ F_۴ = ۹ : \text{فراوانی } \textbf{مطلق} \text{ طبقه چهارم} = \text{فراوانی مطلق طبقه دهک‌دار} \\ P_{C_۴} = f_{C_۴} \times ۱۰۰ = ? : \text{درصد فراوانی نسبی تجمعی طبقه دهک‌دار (چهارم)} \end{cases}$$

خب، با توجه به داده‌های مسئله، برای محاسبه $P_{C_۴}$ باید از فرمول محاسبه دهک ششم استفاده کنیم:

$$\boxed{D_۶ = L_۱ + \frac{\frac{۶N}{۱۰} - F_{C_{i-۱}}}{F_i} \times I}$$

برای استفاده از فرمول بالا، اول باید مقدار L_i و I رو بدست بیاریم:

گام۲) محاسبه فاصله طبقات (I):

$$I = \frac{R}{K} = \frac{X_{\max} - X_{\min}}{K} = \frac{۵۴ - ۱۲}{۷} = \frac{۴۲}{۷} = ۶ \; : \text{فاصله طبقات}$$

(ادامه تو فیش بعد)

گام۳) محاسبه حد پائین واقعی طبقه دهک‌دار (طبقه ششم):

راه اول: برای محاسبه L_i، اول باید جدول فراوانی رو تشکیل بدیم و با توجه به $I = ۶$ و $X_{\min} = ۱۲$ طبقات‌مون رو بنویسیم تا به طبقه چهارم (یعنی طبقه دهک‌دار) برسیم:

CL	$\overbrace{12\text{-}18}^{I=6}$	18-24	24-30	$\overbrace{\boxed{30}\text{-}36}^{\text{طبقه چهارم (طبقه دهک‌دار)}}$
F_i				$F_۴ = ۹$

خب، با توجه به جدول بالا می‌بینیم که:

$L_i = ۳۰$: **حد پائین واقعی طبقه دهک‌دار**

توضیح: چون کلاً $k = ۷$ تا طبقه داریم، پس طبقه وسط که طبق گفته سؤال، طبقه دهک داره، میشه طبقه چهارم.

راه دوم (تستی): *برای رسیدن به حد پائین طبقه چهارم، فقط کافیه به اندازهٔ ۳I به حد پائین طبقه اول (یعنی ۱۲) اضافه کنیم:*

$$L_۴ = L_۱ + ۳I \Rightarrow L_۴ = ۱۲ + ۳(۶) = ۱۲ + ۱۸ \Rightarrow L_۴ = ۳۰$$

$$(۱۲-۱۸) \xrightarrow{I=6} (L_۲ - U_۲) \xrightarrow{I=6} (L_۳ - U_۳) \xrightarrow{I=6} (L_۴ - U_۴),...$$

(ادامه تو پشت فیش)

خب، حالا تازه می‌تونیم فرمول دهک ششم رو بنویسم:

گام۴) استفاده از فرمول $D_۶$:

$$\boxed{D_۶ = L_i + \frac{\frac{۶N}{۱۰} - F_{C_{i-۱}}}{F_i} \times I} \Rightarrow$$

$$۳۲ = ۳۰ + \frac{\frac{۶ \times \cancel{۱۲۰}}{\cancel{۱۰}} - F_{C_{i-۱}}}{۹} \times ۶ \Rightarrow ۳۲ = ۳۰ + \frac{۶ \times ۱۲ - F_{C_{i-۱}}}{\cancel{۹}} \times \cancel{۶} \Rightarrow$$

$$۳۲ = ۳۰ + \frac{(۷۲ - F_{C_{i-۱}}) \times ۲}{۳} \Rightarrow ۳۲ = ۳۰ + \frac{۱۴۴ - ۲F_{C_{i-۱}}}{۳} \times \Rightarrow$$

$$۳۲ - ۳۰ = \frac{۱۴۴ - ۲F_{C_{i-۱}}}{۳} \Rightarrow$$

$$\Rightarrow ۲ = \frac{۱۴۴ - ۲F_{C_{i-۱}}}{۳} \xrightarrow{\text{طرفین وسطین}} ۱۴۴ - ۲F_{C_{i-۱}} = ۶$$

$$\Rightarrow ۱۴۴ - ۶ = ۲F_{C_{i-۱}} \Rightarrow ۱۳۸ = ۲F_{C_{i-۱}} \Rightarrow F_{C_{i-۱}} = \frac{۱۳۸}{۲} \Rightarrow$$

فراوانی تجمعی طبقه قبل از طبقه دهک‌دار : $F_{C_{i-۱}} = F_{C_{۴-۱}} = F_{C_۳} = ۶۹$

(ادامه تو فیش بعد)

خب، ولی ما برای محاسبه $P_{C_۴}$ نیاز $F_{C_۴}$ داریم، یعنی فراوانی تجمعی خود طبقه دهک‌دار:

$$P_{C_۴} = f_{C_۴} \times ۱۰۰ \Rightarrow \frac{F_{C_۴}}{N} \times ۱۰۰$$

گام ۵) محاسبه $F_{C_۴}$:

برای این کار باید فراوانی تجمعی ما قبل، یعنی $F_{C_۳} = ۶۹$ رو با فراوانی مطلق طبقه چهارم ($F_۴ = ۹$) جمع بزنیم:

CL	طبقه سوم ۲۴-۳۰	طبقه چهارم ۳۰-۳۶
F_i		$F_۴ = ۹$
F_{C_i}	$F_{C_۳} = ۶۹$	$F_{C_۴} = ۶۹ + ۹ = ۷۸$

جمع

یعنی:

$$\boxed{F_{C_۴} = F_{C_۳} + F_۴ = ۶۹ + ۹ = ۷۸}$$

گام آخر) محاسبه $P_{C_۴}$:

$$P_{C_۴} = f_{C_۴} \times ۱۰۰ \Rightarrow \frac{F_{C_۴}}{N} \times ۱۰۰ = \frac{۷۸}{\cancel{۱۲۰}} \times \cancel{۱۰۰} \Rightarrow \frac{۷۸}{۱۲} \times ۱۰ \Rightarrow$$

$$= \frac{۲۶}{۴} \times ۱۰ \Rightarrow \frac{۱۳}{۲} \times ۱۰ \Rightarrow \frac{۱۳۰}{۲} = ۶۵ \xrightarrow{\text{پس}} \boxed{P_{C_۴} = ۶۵ \text{ درصد}}$$

مهم:

(مدیریت و حسابداری ۸۹)

متغیرهای پیوسته در جدول زیر گروه‌بندی شده‌اند.

متغیر ۸۰ درصدی داده‌ها کدام است؟

حدود دسته	۱۲-۱۵	۱۵-۱۸	۱۸-۲۱	۲۱-۲۴	۲۴-۲۷	۲۷-۳۰
فراوانی	۵	۱۰	۹	۱۱	۸	۷

۱) ۲۶/۱۲۵

۲) ۲۵/۶۲۵

۳) ۲۵/۸۷۵

۴) ۲۶/۲۲۵

پاسخ: گزینه ۳

توجه: منظور از **متغیر ۸۰ درصدی** داده‌ها، همان **صدک هشتادم** $(P_{۸۰})$ است.

گام ۱) محاسبه فراوانی تجمعی طبقات:

CL	۱۲-۱۵	۱۵-۱۸	۱۸-۲۱	۲۱-۲۴	۲۴-۲۷	۲۷-۳۰
F_i	۵	۱۰	۹	۱۱	۸	۷
F_{C_i}	۵	۱۵	۲۴	۳۵	۲۳	$N = ۵۰$

گام ۲) پیدا کردن محل صدک هشتادم:

$$C_{P_{۸۰}} = \frac{۸۰N}{۱۰۰} \xrightarrow{N=۵۰} \frac{۸۰ \times \cancel{۵۰}}{\cancel{۱۰۰}} = \frac{۸۰}{۲} = ۴۰$$: **محل صدک ۸۰اُم**

یعنی ۴۰اُمین داده، صدک هشتادمه.

پیدا کردن طبقه صدک‌دار:

اولین طبقه‌ای که فراوانی تجمعی‌اَش **بزرگتر یا مساوی** $\frac{۸۰N}{۱۰۰} = ۴۰$ باشه $(F_{C_i} \geq ۴۰)$، طبقه (۲۷-۲۴) است:

$$\xrightarrow{\text{طبقه پنجم}} F_{C_۵} = ۴۳ \geq ۴۰ \xrightarrow{\text{پس}} \boxed{(۲۷-۲۴) : \text{طبقه صدک‌دار}}$$

(ادامه تو فیش بعد)

گام ۳) محاسبه مقدار صدک ۸۰اُم:

$$\boxed{\textbf{مقدار صدک هشتادم}: P_{۸۰} = L_{۵} + \left(\frac{\frac{۸۰N}{۱۰۰} - F_{C_{i-۱}}}{F_i}\right) \times I} \Rightarrow$$

$$P_{۸۰} = ۲۴ + \left(\frac{۴۰ - ۳۵}{۸}\right) \times ۳ \Rightarrow$$

$$P_{۸۰} = ۲۴ + \frac{۵}{۸} \times ۳ \Rightarrow$$

$$P_{۸۰} = ۲۴ + \frac{۵ \times ۳}{۸} \Rightarrow$$

$$P_{۸۰} = ۲۴ + \frac{۱۵}{۸} \Rightarrow$$

$$P_{۸۰} = ۲۴ + ۱/۸۷۵ \Rightarrow$$

$$\boxed{P_{۸۰} = ۲۵/۸۷۵}$$

یعنی **متغیر ۸۰ درصدی داده‌ها** مساوی ۲۵/۸۷۵ است.

(مهم):

با استفاده از جدول زیر، دهک هفتم را محاسبه کنید.

C-L	F_i
۱۰–۲۰	۱۵
۲۰–۳۰	۳۰
۳۰–۴۰	۲۵
۴۰–۵۰	۲۰
۵۰–۶۰	۱۰
	$N = ۱۰۰$

۱) ۴۱ ۲) ۴۵

۳) ۴۰ ۴) ۴۴

توجه: از حل این سؤال به چه نکته‌ای پی بردین؟

آمار و کاربرد آن در مدیریت، عادل آذر و منصور مؤمنی، جلد ۱، ص ۱۰۷ و ۱۰۸

گام ۱) محاسبه فراوانی تجمعی طبقات (در جدول زیر):

C-L	F_i	F_{C_i}
۱۰-۲۰	۱۵	۱۵
۲۰-۳۰	۳۰	$(۳۰+۱۵)=۴۵$
۳۰-۴۰	۲۵	$(۲۵+۴۵)=\boxed{۷۰} \rightarrow$ طبقه دهک‌دار (۴۰-۳۰)
۴۰-۵۰	۲۰	$(۲۰+۷۰)=۹۰$
۵۰-۶۰	۱۰	$(۹۰+۱۰)=۱۰۰$
	$N=۱۰۰$	

گام ۲) یافتن محل دهک هفتم (طبقه دهک‌دار):

اولین طبقه از بالا به پایین که فراوانی تجمعیش بزرگتر یا مساوی $F_{C_i} \geq \frac{aN}{۱۰}$ باشد، طبقه دهک‌دار است:

محل دهک هفتم: $\frac{۷N}{۱۰} \xrightarrow{N=۱۰۰} \frac{۷\times\cancel{۱۰۰}}{\cancel{۱۰}}=۷\times۱۰=۷۰ \rightarrow$ دارای ۷۰ اُم

پیدا کردن طبقه دهک‌دار (روش اول): اولین طبقه‌ای که فراوانی تجمعیش بزرگتر یا مساوی $\frac{۷N}{۱۰}=۷۰$ باشد، طبقه سومه (۴۰-۳۰)، پس:

$\xrightarrow{\text{طبقه سوم}} F_C=۷۰\geq۷۰ \xrightarrow{\text{پس}}$ **طبقه دهک‌دار** : (۴۰-۳۰)

(ادامه تو فیش بعد)

پیدا کردن طبقه دهک‌دار (روش دوم): *با نگاه به جدول می‌فهمیم که از دادهٔ ۴۶ اُم تا ۷۰ اُم تو طبقه (۴۰- ۳۰) قرار دارن، پس دادهٔ ۷۰ اُم (یعنی دهک هفتم) هم حتماً تو این طبقه (۴۰-۳۰) قرار می‌گیره.*

توجه: چون طبقات ما **پیوسته** هستن، پس دیگه لازم نیست تا طبقه چارک‌دار (۴۰-۳۰) رو پیوسته کنیم.

مقدار دهک هفتم : $D_۷ \approx L_i + \dfrac{\frac{aN}{۱۰} - F_{C_{i-۱}}}{F_i} \times I \Rightarrow$

$$D_۷ \approx ۳۰ + \frac{۷۰ - ۴۵}{۲۵} \times ۱۰ \Rightarrow D_۷ \approx ۳۰ + \frac{۲۵}{۲۵} \times ۱۰ \Rightarrow$$

$$D_۷ \approx ۳۰ + (۱ \times ۱۰) \Rightarrow D_۷ \approx ۳۰ + ۱۰ \Rightarrow$$

دهک هفتم : $D_۷ = ۴۰$

نکته مهم: در این سؤال محل چندک ($C_{D_۷} = \frac{aN}{۱۰} =$ ۷۰ اُمی) **دقیقاً در ستون فراوانی تجمعی دیده می‌شود** ($F_{C_۳} = ۷۰$)، در این شرایط می‌توان **آن طبقه** (طبقهٔ چندک‌دار) **یا طبقه پایین‌تر از آن (بعد از آن)** را به عنوان طبقه چندک‌دار در نظر گرفت و نتیجه هر دو حالت، یکسان خواهد بود.

(ادامه تو پشت فیش)

مثلاً اگر در این سؤال، طبقه **پایین‌تر**، یعنی طبقه چهارم را به عنوان طبقه چندک‌دار در نظر بگیریم، باز هم به $D_۷ = ۴۰$ می‌رسیم: *(۴۰–۵۰)؛ طبقه چندک‌دار*

$$\boxed{\textbf{مقدار دهک هفتم}: D_۷ \approx L_i + \frac{\frac{aN}{۱۰} - F_{C_{i-۱}}}{F_i} \times I \Rightarrow}$$

$$D_۷ \approx ۴۰ + (\frac{\cancel{۷۰-۷۰}}{۲۰}) \times ۱۰ \Rightarrow D_۷ \approx ۴۰ + (\cancel{۰ \times ۱۰}) \Rightarrow$$

$$D_۷ \approx ۴۰ + ۰ \Rightarrow \boxed{D_۷ \approx ۴۰}$$

نتیجه: همون‌طور که می‌بینین جواب هر دو روش دقیقاً مثل همه.

راه‌حل و نکته تستی: اگه تو سؤال، **محل** چندک (یعنی $\frac{aN}{۴}$، $\frac{aN}{۱۰}$ یا $\frac{aN}{۱۰۰}$) رو **دقیقاً** تو بین مقادیر فراوانی **تجمعی** ببینیم (مثل این سؤال که محل و دهک هفتم، یعنی دادهٔ ۷۰ اُم، **دقیقاً** تو ستون فراوانی **تجمعی** دیده میشه: $F_{C_۳} = ۷۰$)، تو چنین حالتی **بدون نیاز به هیچ محاسبه‌ای** میتونیم بگیم که:

چندک مورد نظر ما برابر با حد بالای طبقه چندک داره.

مثلاً تو این سؤال، می‌تونستیم *بدون هیچ محاسبه‌ای*، خیلی سریع بگیم که دهک هفتم مساوی **حد بالای** طبقه سومه (۴۰-۳۰)، یعنی: $D_{۷} = ۴۰$

اما یادمون باشه که: فقط موقعی می‌تونیم از این راه حل تستی استفاده کنیم که طبقات جدول ما **از همون اول پیوسته باشن**، یعنی مثل همین سؤال.

(مهم):

«نحوهٔ محاسبه چندک‌ها در داده‌های نوع سوم (حجم زیاد و تنوع زیاد)

با داشتن فراوانی نسبی (f_i)»

مثال: *با توجه به جدول طبقه‌بندی شده زیر، مقدار دهک ششم کدام است؟ (آن را تفسیر کنید).*

a	۰–۶	۶–۱۲	۱۲–۱۸	۱۸–۲۴
f_i	۰/۳	۰/۱	۰/۴	۰/۲

۱) ۱۲/۵

۲) ۱۵

۳) ۱۱

۴) ۶

آمار و احتمال، آموزشگاه پارسه، محسن طورانی، ص ۴۲

«محاسبه چندک‌ها در داده‌های نوع سوم (طبقه‌بندی) با داشتن (f_i)»

گام ۱) محاسبه فراوانی تجمعی نسبی (f_i)

گام ۲) یافتن محل چندک (طبقه چندک‌دار):

اولین طبقه‌ای که فراوانی تجمعی نسبیش:

۱) برای **چارک** aاُم بزرگتر یا مساوی $\frac{a}{۴}$ باشد ($f_{C_i} \geq \frac{a}{۴}$)

مثلاً: برای چارک سوم باید $f_{C_i} \geq \frac{۳}{۴} = ۰/۷۵$ باشد.

۲) برای **دهک** aاُم بزرگتر یا مساوی $\frac{a}{۱۰}$ باشد ($f_{C_i} \geq \frac{a}{۱۰}$)

مثلاً: برای دهک ششم ($a = ۶$)، باید داشته باشیم: $f_{C_i} \geq \frac{۶}{۱۰} = ۰/۶$

۳) برای **صدک** aاُم بزرگتر یا مساوی $\frac{a}{۱۰۰}$ باشد ($f_{C_i} \geq \frac{a}{۱۰۰}$)

مثلاً: برای صدک بیستم ($a = ۲۰$)، باید: $f_{C_i} \geq \frac{۲۰}{۱۰۰} = ۰/۲۰$

گام ۳) پیوسته کردن طبقات در صورت گسسته بودن

گام ۴) یافتن مقدار تقریبی چندک aاُم (a: شماره چندک) از روابط زیر:

(ادامه تو فیش بعد)

توجه: با دقت در روابط زیر متوجه می‌شیم که اگر بجای فراوانی مطلق (F_i) فراوانی **نسبی** (f_i) را داشته باشیم، اون‌وقت برای محاسبه چندک‌ها، به همون روش قبلی عمل می‌کنیم، فقط با این تفاوت که در فرمول‌ها **به جای** N **از عدد یک** استفاده می‌کنیم، زیرا در واقع در اینجا N به مفهوم مجموع فراوانی‌های نسبی است که برابر با **یک** است:

$$Q_a \approx L_i + \frac{\frac{aN}{۴} - f_{C_{i-۱}}}{f_i} \times I \xrightarrow{N=۱} \boxed{Q_a \approx L_i + \frac{\frac{a}{۴} - f_{C_{i-۱}}}{f_i} \times I}$$

$$D_a \approx L_i + \frac{\frac{aN}{۱۰} - f_{C_{i-۱}}}{f_i} \times I \xrightarrow{N=۱} \boxed{D_a \approx L_i + \frac{\frac{a}{۱۰} - f_{C_{i-۱}}}{f_i} \times I}$$

$$P_a \approx L_i + \frac{\frac{aN}{۱۰۰} - f_{C_{i-۱}}}{f_i} \times I \xrightarrow{N=۱} \boxed{P_a \approx L_i + \frac{\frac{a}{۱۰۰} - f_{C_{i-۱}}}{f_i} \times I}$$

حل مثال: مقدار دهک ششم در داده‌های جدول زیر کدام است؟

C-L	۰-۶	۶-۱۲	۱۲-۱۸	۱۸-۲۴
f_i	۰/۳	۰/۱	۰/۴	۰/۲

پاسخ: گزینه ۲

گام ۱) محاسبه f_{C_i} (فراوانی تجمعی نسبی) در جدول:

CL	0-6	6-12	12-18	18-24	
f_i	۰/۳	۰/۱	۰/۴	۰/۲	$\sum f_i = N = 1$
f_{C_i}	0/3	0/4	0/8	1	

گام ۲) یافتن محل دهک ششم (طبقه دهک‌دار):

اولین طبقه‌ای که در آن فراوانی تجمعی نسبی بزرگتر یا مساوی $\frac{a}{۱۰} = \frac{۶}{۱۰} = ۰/۶$ باشد ($f_{C_i} \geq ۰/۶$). طبقه (۱۸-۱۲) است، پس این طبقه، طبقه دهک‌دار ماست:

$$\xrightarrow{\text{طبقه دوم}} f_{C_۳} = ۰/۸ \geq ۰/۶ \xrightarrow{\text{پس}} \boxed{\text{(۱۸-۱۲): طبقه دهک‌دار}}$$

توجه: طبقات ما **پیوسته** هستن، پس دیگه نمی‌خواد طبقه (۱۸-۱۲) رو پیوسته کنیم، چون خودش پیوسته هست، در نتیجه:

$\boxed{I = ۱۸ - ۱۲ = ۶}$, $\boxed{L_i = ۱۲ : \text{حد پایین واقعی طبقه دهک‌دار}}$

گام ۳) محاسبه مقدار تقریبی دهک ششم از رابطه زیر:

(ادامه تو فیش بعد)

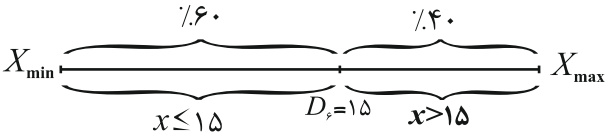
مقدار دهک ششم: $D_۶ \approx L_i + \dfrac{\frac{a}{۱۰} - f_{C_{i-۱}}}{f_i} \times I \Rightarrow$

$$D_۶ \approx ۱۲ + \frac{۰/۶ - ۰/۴}{۰/۴} \times ۶ \Rightarrow D_۶ \approx ۱۲ + \frac{۰/۲}{۰/۴} \times ۶ \Rightarrow$$

$$D_۶ \approx ۱۲ + \frac{۲}{۴} \times ۶ \Rightarrow D_۶ \approx ۱۲ + (\frac{۱}{۲} \times ۶) \Rightarrow$$

$$D_۶ \approx ۱۲ + ۳ \Rightarrow \boxed{D_۶ \approx ۱۵}$$

مفهومِ $D_۶ \approx ۱۵$: یعنی $\frac{۶}{۱۰} = ٪۶۰$ مشاهدات **کوچکتر یا مساوی** $x = ۱۵$ و $\frac{۴}{۱۰} = ٪۴۰$ مشاهدات باقیمانده هم **بزرگتر** از $x = ۱۵$ هستند:



نکته تستی: *مشابه همون نکته‌ای که قبلاً در مورد طبقات با فراوانی* **مطلق** *(F_i) گفتیم، تو اینجا هم در مورد فراوانی نسبی (f_i) می‌تونیم بگیم:*

(ادامه تو پشت فیش)

اگه طبقات جدول ما **از همون اول، خودشون پیوسته باشن،** *اون وقت اگه محل چندک ما دقیقاً توسط فراوانی تجمعی نسبی (F_i) وجود داشته باشد، تو این حالت،خیلی راحت و بدون نیاز به هیچ محاسبه‌ای می‌تونیم بگیم که:*

چندک مورد نظر ما = **حد بالای** *طبقه چندک داره*

مثال۱: اگه تو همین سؤال که طبقاتش، **از همون اول پیوسته هستن،** مقدار دهک سوم یا صدک سی‌اُم رو از ما می‌خواستن، اون وقت چون مقدار $f_C = ۰/۳$ دقیقاً توسط f_{C_i} مربوط به طبقه (۶-۰) دیده میشه، پس خیلی راحت می‌تونیم بگیم که:

$D_۳ = P_{۳۰} = ۶ \Rightarrow$ **حد بالای** *طبقه (۰-۶)*

مثال۲: اگه از ما مقدار دهک چهارم یا صدک چهلم رو بخوان، چون $f_C = ۰/۴$ تو سطر f_{C_i} دیده میشه، پس **حد بالای** این طبقه (۱۲-۶) همون چندک مورد نظر ما خواهد بود:

$D_۴ = P_{۴۰} = ۱۲ \Rightarrow$ **حد بالای** *طبقه (۶-۱۲)*

مثال۳: و اگر از ما مقدار دهک هشتم یا صدک هشتادم رو می‌خواستن، از اونجایی که تو ستون مربوطه به طبقه (۱۸-۱۲) می‌تونیم **عیناً** $f_C = ۰/۸$ رو ببینیم، پس می‌تونیم بگیم که:

$D_۸ = P_{۸۰} = ۱۸ \Rightarrow$ **حد بالای** *طبقه (۱۲-۱۸)*

بسیار مهم:

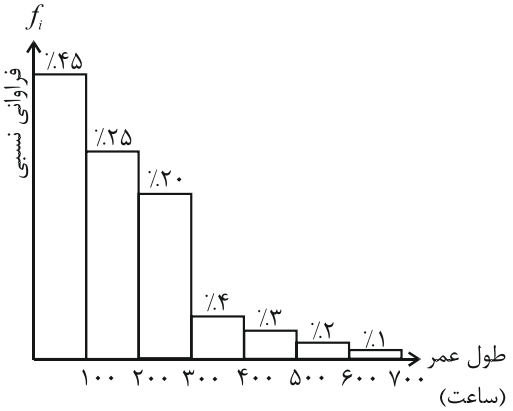
(محیط زیست – سراسری ۸۷)

توزیع طول عمر ۱۰۰۰ قطعه ترانزیستوری که در یک کارخانه تولید شده است، به شکل زیر است.

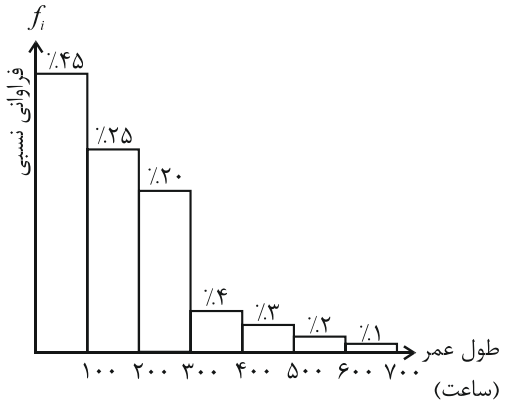
اگر بدانیم ۲۰٪ از قطعه‌ها از درجه کیفیت کمتری برخوردار می‌باشند و دارای طول عمر پایین هستند، پس این تعداد کمتر از چه مدت عمر می‌کنند؟

۱) ۲۰ ۲) ۴۴/۴

۳) ۵۰/۵ ۴) ۷۰



پاسخ: گزینه ۲



گام ۱) نوشتن جدول توزیع فراوانی و محاسبه فراوانی تجمعی نسبی :

طول عمر (ساعت)	۰–۱۰۰	۱۰۰–۲۰۰	۲۰۰–۳۰۰	۳۰۰–۴۰۰	۴۰۰–۵۰۰	۵۰۰–۶۰۰	۶۰۰–۷۰۰
فراوانی نسبی	٪۴۵	٪۲۵	٪۲۰	٪۴	٪۳	٪۲	٪۱
فراوانی نسبی تجمعی f_{C_i}	٪۴۵	٪۷۰	٪۹۰	٪۹۴	٪۹۷	٪۹۹	٪۱۰۰

گام ۲) فهمیدن صورت مسئله:

می‌دونیم که صدک بیستم، مقداری از X است که ۲۰٪ از داده‌ها کمتر یا مساوی اون هستن. (ادامه تو فیش بعد)

خب، این سؤال، از ما خواسته که طـول عمـری (xای) رو پیـدا کنـیم کـه ۲۰٪ از قطعه‌ها **کمتر یا مساوی** اون، عمر می‌کنن. یعنی در واقع به زبون بی زبونی از ما خواسته تا **مقدار صدک بیستم** رو حساب کنیم.

برای بدست آوردن صدک بیستم $(P_{۲۰})$ به صورت زیر عمل می‌کنیم:

گام ۲) یافتن محل صدک ۲۰ اُم:

اولین طبقه‌ای که فراوانی تجمعی نسـبیش، **بزرگتـر یـا مسـاوی** ۰/۲ یـا ۲۰٪ باشد، طبقه اول (۱۰۰-۰) است. پس صدک بیستم در طبقه اول قرار دارد.

$$\xrightarrow{\text{در طبقه اول}} f_C = ۰/۴۵ \geq ۰/۲۰ \xrightarrow{\text{پس}} \boxed{(۰-۱۰۰) : \text{طبقه صدک‌دار}}$$

توجه کنین که:

تو این سؤال، طبقات ما **پیوسـته** هستن، پـس دیگـه نیـازی بـه پیوسـته کردن طبقه (۱۰۰-۰) نداریم، چـون خـودش از قبـل **پیوســته** اسـت. در نتیجه می‌تونیم بگیم که:

$\boxed{I = ۱۰۰ - ۰ = ۱۰۰}$ و $\boxed{L_i = ۰ : \text{حد پایین واقعی طبقه صدک‌دار}}$

(ادامه تو پشت فیش)

گام ۴) محاسبه مقدار صدک:

$$\boxed{P_{۲۰} = L_i + \frac{۰/۲ - f_{C_{i-۱}}}{f_i} \times I}$$

$$P_{۲۰} = ۰ + \frac{۰/۲ - ۰}{۰/۴۵} \times ۱۰۰ \Rightarrow$$

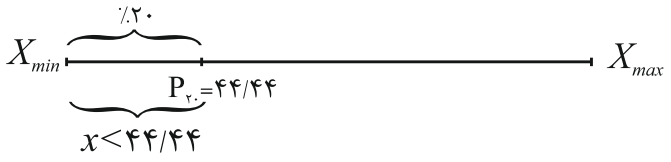
$$P_{۲۰} = \frac{۰/۲}{۰/۴۵} \times ۱۰۰ \Rightarrow$$

صورت و مخرج رو در ۱۰۰ ضرب می‌کنیم تا

$$P_{۲۰} = \frac{۰/۲ \times ۱۰۰}{۰/۴۵} = \frac{۲۰}{۰/۴۵} \xrightarrow{\text{اعشار ۰/۴۵ از بین بره}}$$

$$\Rightarrow \frac{۲۰۰۰}{۴۵} = ۴۴/۴۴ \xrightarrow{\text{پس}} \boxed{\text{ساعت } ۴۴/۴۴ = P_{۲۰} : \textbf{صدک بیستم}}$$

یعنی ۲۰٪ از قطعه‌ها ***کمتر از*** ۴۴/۴۴ ساعت عمر می‌کنن:



(مهم):

اصلی‌ترین، مهمترین و پرکاربردترین شاخص مرکزی که نشان دهندهٔ **نقطهٔ تعادل و مرکز ثقل** جامعه است، .....(۱)..... یک جامعه است.

در حقیقت اگر مشاهدات جامعه (x_i ها) را روی یک محور به صورت منظم ردیف کنیم، آنگاه .....(۲).....، **نقطه تعادل و ثقل** جامعه خواهد بود.

این محور همانند یک آلاکلنگ است که .....(۳).....، نقطه تعادل آن است، به گونه‌ای که جمع جبری گشتاور (یا تفاضلِ) داده‌ها نسبت به .....(۴).....، برابر .....(۵)..... خواهد شد.

آمار و احتمال، آموزشگاه پارسه، محسن طورانی،ص ۴۳

۱) میانگین (حسابی) **۲)** میانگین (حسابی)

۳) میانگین (حسابی) **۴)** میانگین (حسابی)

۵) صفر

اصلی‌ترین، مهمترین و پرکاربردترین شاخص مرکزی که نشان دهندهٔ **نقطهٔ تعادل و مرکز ثقل** جامعه است، **میانگین (حسابی)** یک جامعه است.

در حقیقت اگر مشاهدات یک جامعه (یعنی x_i ها) را به صورت منظم روی یک محور ردیف کنیم، آنگاه **میانگین (حسابی)**، **نقطه تعادل و ثقل** جامعه خواهد بود.

این محور همانند الاکلنگ است که *میانگین (حسابی)*، *نقطه تعادل* آن است، به گونه‌ای که جمع جبری گشتاور (یا تفاضلِ) داده‌ها نسبت به **میانگین (حسابی)**، برابر **صفر** خواهد شد.

مثال: $x_i : ۱,۲,۳ \Rightarrow$ **میانگین** $= \frac{۱+۲+۳}{۳} = \frac{۶}{۳} \Rightarrow \boxed{\mu = ۲}$

گشتاور یا تفاضل داده‌ها از میانگین $= (x_i - \mu) = (۱-۲) + (۲-۲) + (۳-۲)$

$$\Rightarrow -۱ + ۰ + ۱ = ۰$$

به عبارت دیگر می‌توان گفت:

$$\boxed{\sum (x_i - \mu) = \sum (x_i - ۲) = ۰}$$