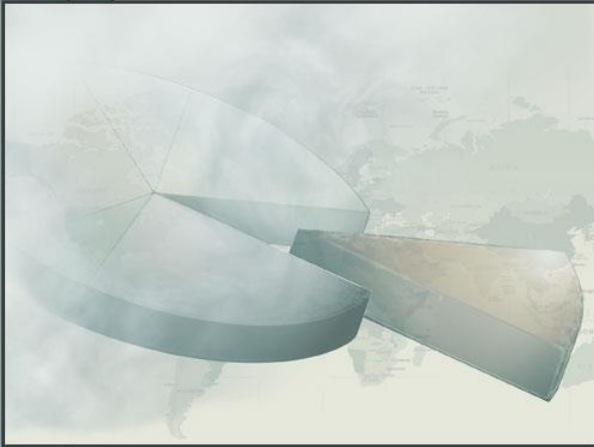


به نام دادار یکتا

آمار توصیفی



تعریف آمار:

جمع آوری

خلاصه کردن و طبقه بندی

تجزیه و تحلیل اطلاعات

کاربردها:

توصیف وضع کنونی

مقایسه وضع کنونی با گذشته

مقایسه وضع خود با دیگران

پیش بینی وضع آینده

سنجش ارتباطات و اثرات

یافتن نقاط قوت و ضعف

فهرست مطالب:

■ آمار توصیفی و استنباطی

■ جدول های آماری

■ نمودار های آماری

■ شاخص های مرکزی

■ شاخص های پراکندگی

■ انواع نمونه

آمار توصیفی و استنباطی:



خلاصه کردن و توصیف خصوصیات مهم جامعه یا نمونه مورد بررسی



تعمیم ها، استنباط ها و نتیجه گیری ها بر اساس نمونه مورد مطالعه برای جامعه مورد نظر

جمعیت و نمونه:

جامعه:

مجموعه کل واحدهای مورد نظر

مثال: مراجعین روزانه یک کتابخانه، دانشجویان عضو کتابخانه، کارکنان کتابخانه های دانشگاه

نمونه:

صرف هزینه، وقت و نیروی انسانی زیاد برای دسترسی و مطالعه کل جامعه و امکان پذیر نبودن آن در برخی موارد: اکتفا به قسمتی از جمعیت آماری (نمونه)

مثال: برای بررسی نظرات دانشجویان عضو کتابخانه، انتخاب مثلا ۱۵۰ نفر از بین این جمعیت به طور تصادفی انتخاب می شوند.

متغیرها:

متغیر:

خصوصیت مورد مطالعه، از یک واحد جامعه به واحد دیگر تغییر می‌کند و قابل اندازه‌گیری است.

انواع متغیرها:

متغیرهای کیفی (گروهی)، نظیر جنس، شغل، رضایت شغلی و ... که شامل چند گروه یا طبقه می‌باشند.

- متغیرهای کیفی اسمی: صرفاً شناسایی (جنس، نژاد، رشته تحصیلی)
- متغیرهای کیفی ترتیبی: شناسایی و بیان برتری (درجه شدت بیماری، مهارت کارکنان)

متغیرهای کمی (عددی): قابل اندازه‌گیری با اعداد و ارقام مانند تعداد کتب موجود در یک کتابخانه، سن، تعداد اعضاء

خلاصه کردن داده‌ها

داده‌های جمع آوری شده که انبوهی عدد است و هیچ نوع پردازشی روی آنها انجام نشده است.



شیوه‌های خلاصه کردن داده‌ها

1. تنظیم جداول
2. خلاصه کردن و توضیح داده‌ها به وسیله رسم نمودارها.
3. خلاصه کردن داده‌ها در چند شاخص توصیفی که بیانگر ویژگی‌های کلی داده باشد.

جدول‌های آماری:

جدول فراوانی: متداول‌ترین جدول در آمار

• **فراوانی:** هرگاه n داده از k نوع، به ترتیب با تعدادهای f_1, f_2, \dots, f_k تشکیل شده باشند، آنگاه f_i فراوانی x_i می‌گوییم.

مثال: تعداد کتب روزانه به امانت برده شده از کتابخانه

x_i	f_i	
2	1	۸ ۶ ۵ ۵ ۳ ۴ ۳ ۶ ۶ ۷
3	4	۲ ۸ ۷ ۶ ۵ ۶ ۶ ۵ ۵ ۶
4	3	۳ ۵ ۵ ۸ ۵ ۷ ۴ ۸ ۳ ۴
5	8	
6	7	
7	3	
8	4	
کل	30	

جدولهای آماری:

فراوانی نسبی: نسبت واحدهای یک گروه به کل واحدها

$$r_i = \frac{f_i}{n}$$

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

$$0 \leq r_i \leq 1$$

$$\frac{1}{30} = 0.033$$

x_i	f_i	r_i
2	1	0.033
3	4	0.133
4	3	0.100
5	8	0.267
6	7	0.233
7	3	0.100
8	4	0.133
کل	30	۱/۰۰۰

$$\frac{4}{30} = 0.133$$

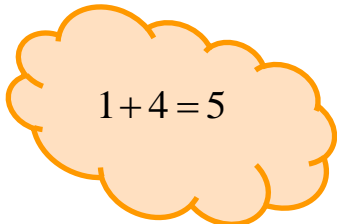
جدول‌های آماری:

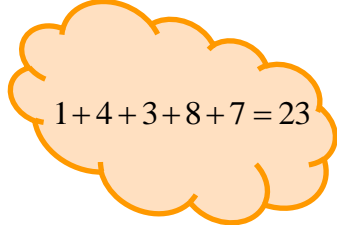
فراوانی تجمعی: تعداد واحدهای یک گروه و گروه‌های ماقبل آن
مثال: تعداد دانشجویان کارشناسی ارشد و پایین‌تر از آن

$$F_i = \sum_{j=1}^i f_j$$

$$f_i = F_1 \leq F_1 \leq \dots \leq F_k = n$$

x_i	f_i	F_i
2	1	1
3	4	5
4	3	8
5	8	16
6	7	23
7	3	26
8	4	30


$$1+4=5$$

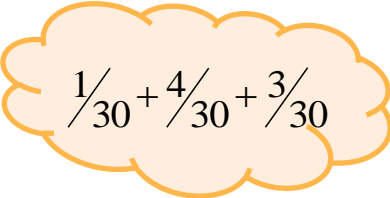

$$1+4+3+8+7=23$$

جدول‌های آماری:

فراوانی نسبی تجمعی: نسبت واحدهای یک گروه و گروه‌های ماقبل آن به کل
مثال: نسبت دانشجویان کارشناسی ارشد و پایین تر از آن

$$R_i = \sum_{j=1}^i r_j$$

x_i	r_i	R_i
2	$\frac{1}{30}$	$\frac{1}{30}$
3	$\frac{4}{30}$	$\frac{5}{30}$
4	$\frac{3}{30}$	$\frac{8}{30}$
5	$\frac{8}{30}$	$\frac{16}{30}$
6	$\frac{7}{30}$	$\frac{23}{30}$
7	$\frac{3}{30}$	$\frac{26}{30}$
8	$\frac{4}{30}$	$\frac{30}{30}$


$$\frac{1}{30} + \frac{4}{30} + \frac{3}{30}$$

نمودارهای آماری:

تجسم عینی اطلاعات نهفته در داده‌ها

ستونی (میل‌های یا نرده‌ای)

دایره‌ای (لوله‌ای)

مسطیلی (هیستوگرام)

چندبر فراوانی (چندضلعی)

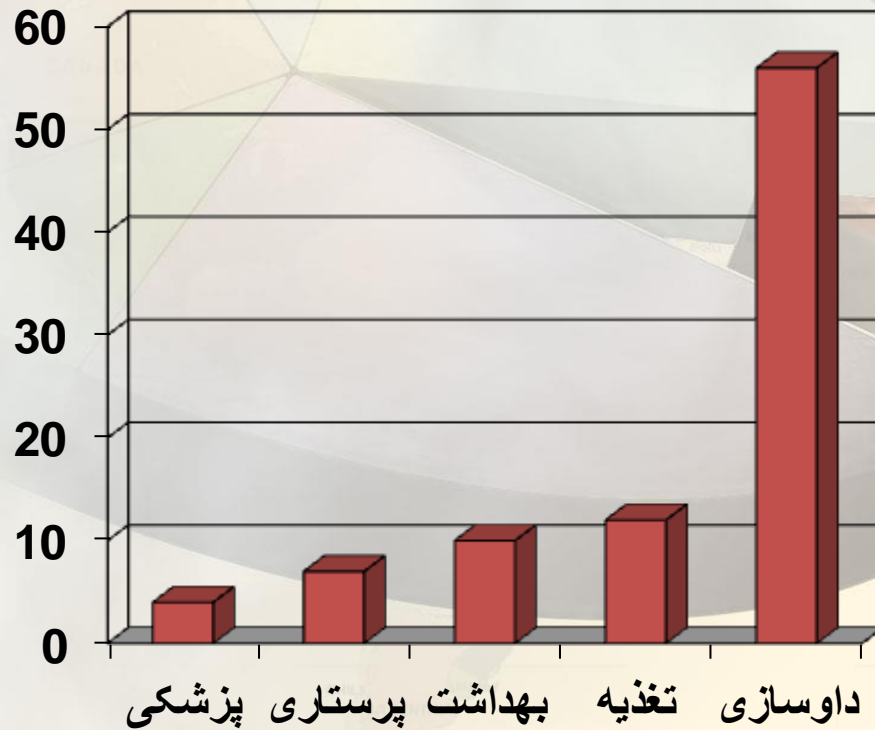
چندبر فراوانی تجمعی

نمودار شاخه و برگ

نمودارهای آماری:

نمودار میله‌ای

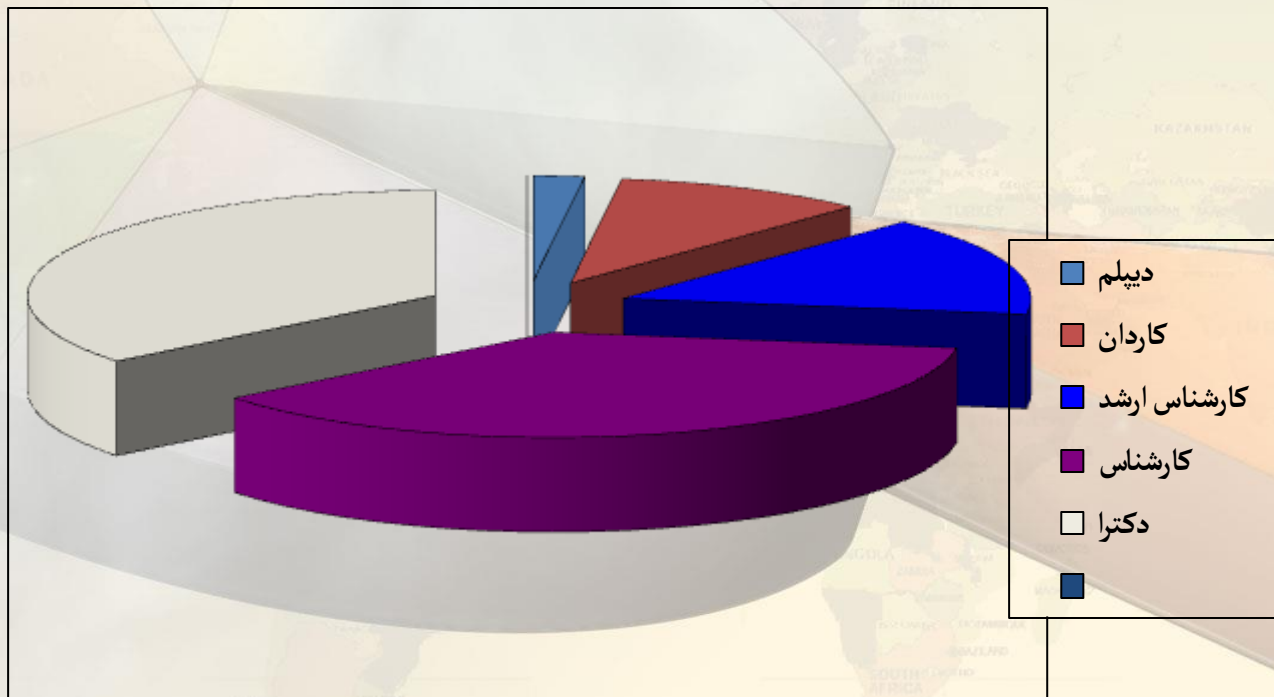
مناسب برای متغیرهای کیفی یا گسسته
نمایش فراوانی هر گروه با یک میله



نمودارهای آماری:

نمودار دایره‌ای

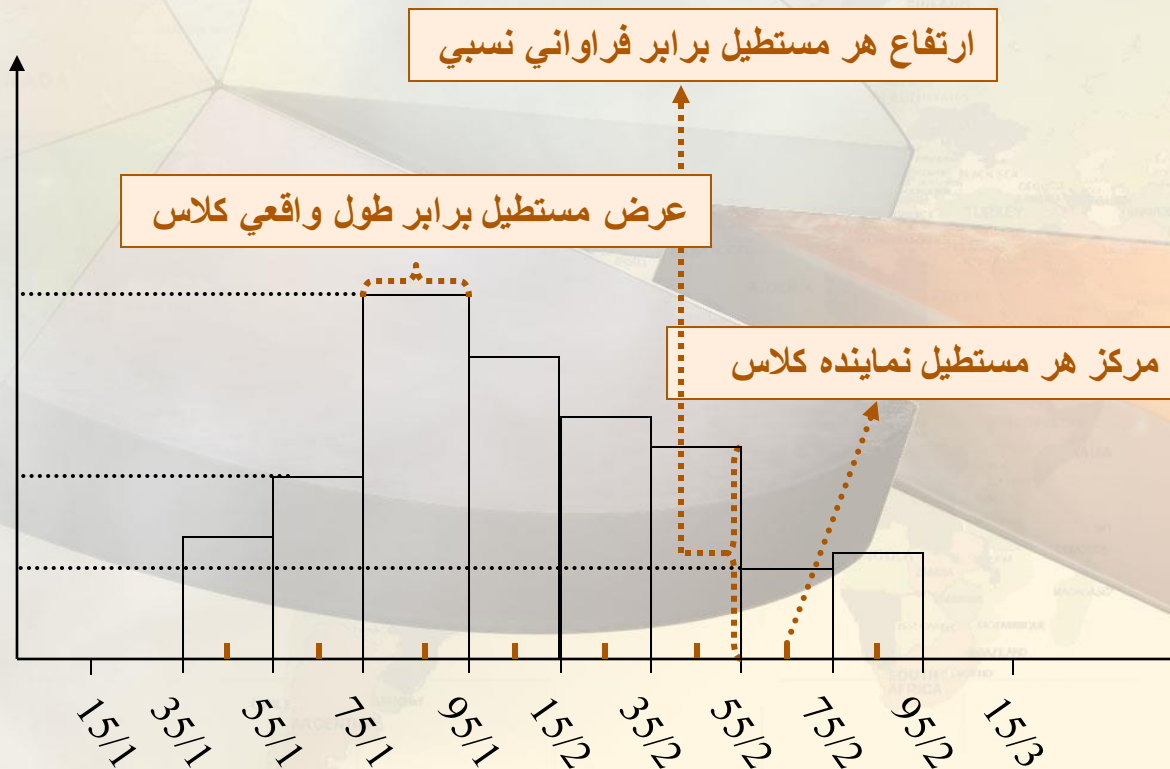
مناسب برای متغیرهای کیفی یا گسسته
نمایش فراوانی هر گروه با بخشی از سطح دایره



نمودارهای آماری:

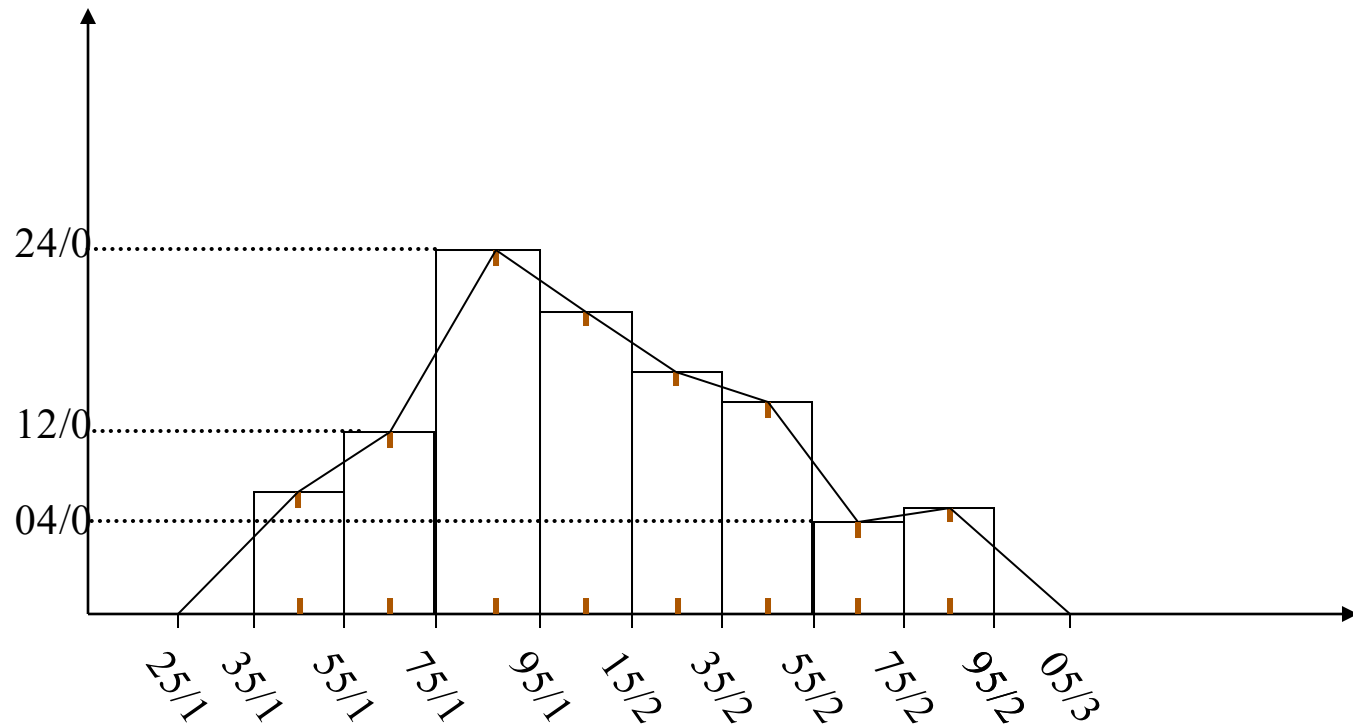
نمودار هیستوگرام

مناسب برای متغیرهای کمی پیوسته
همانند نمودار میله‌ای است با این تفاوت که به جای میله‌ها، از مستطیل برای نمایش فراوانی هر دسته استفاده می‌شود.



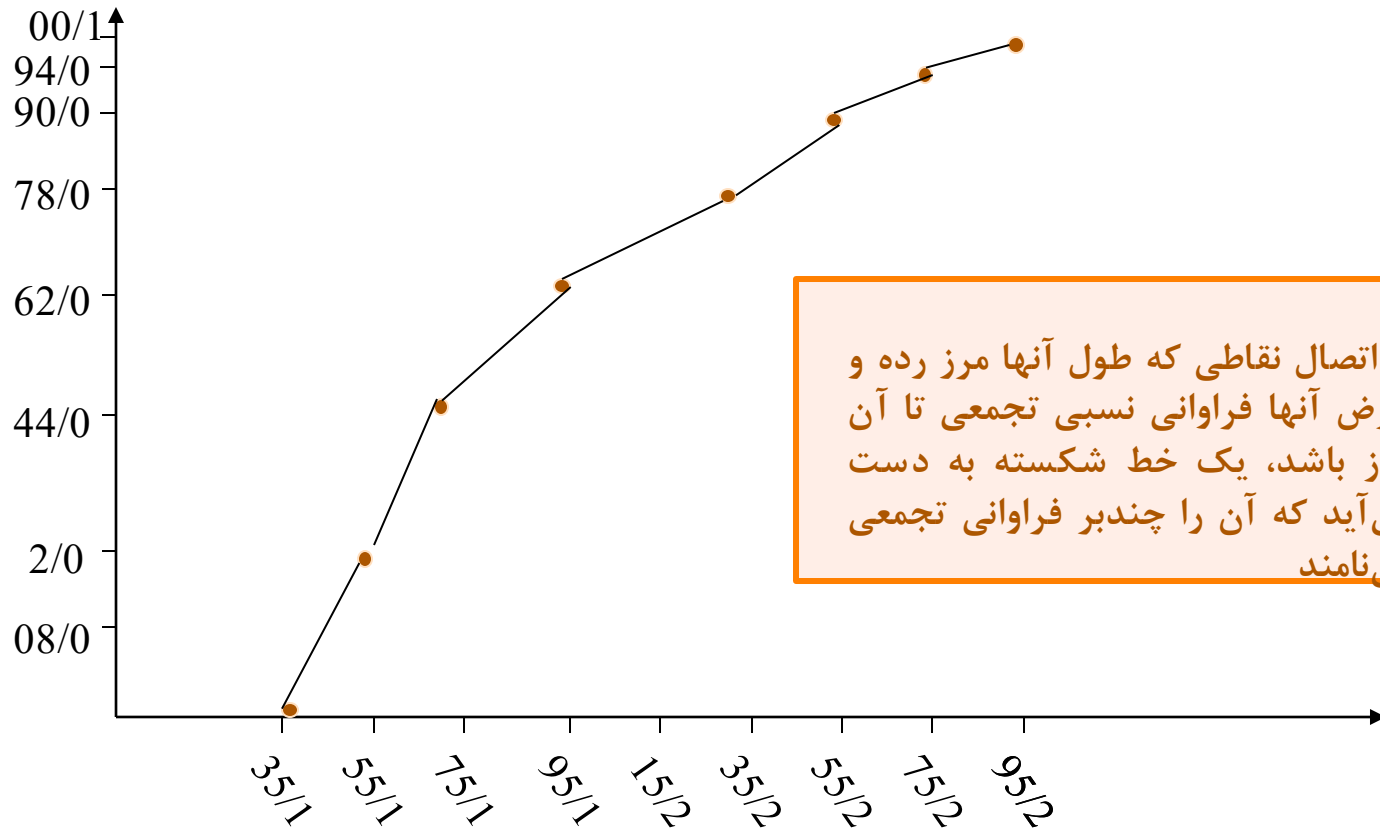
نمودارهای آماری:

چندبر فراوانی



نمودارهای آماری:

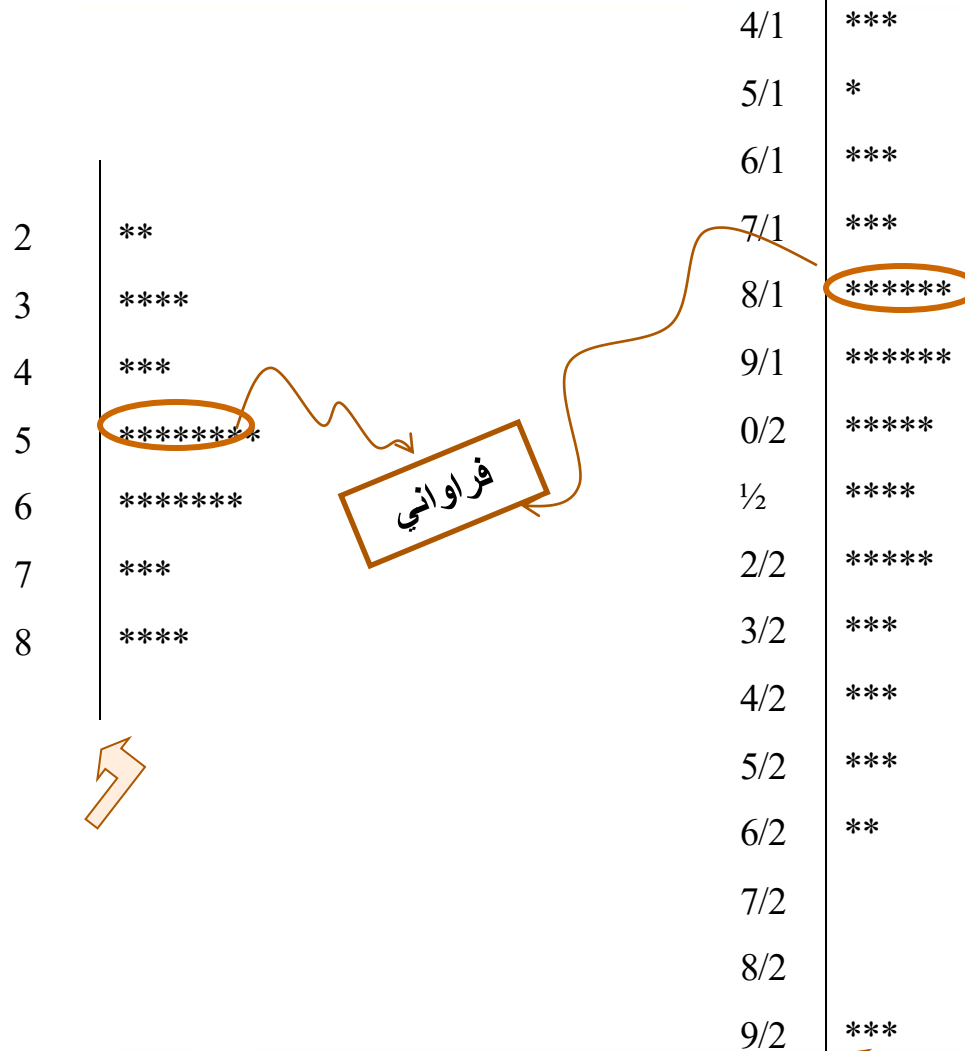
چندبر فراوانی تجمعی



از اتصال نقاطی که طول آنها مرز رده و عرض آنها فراوانی نسبی تجمعی تا آن مرز باشد، یک خط شکسته به دست می آید که آن را چندبر فراوانی تجمعی می نامند

نمودارهای آماری:

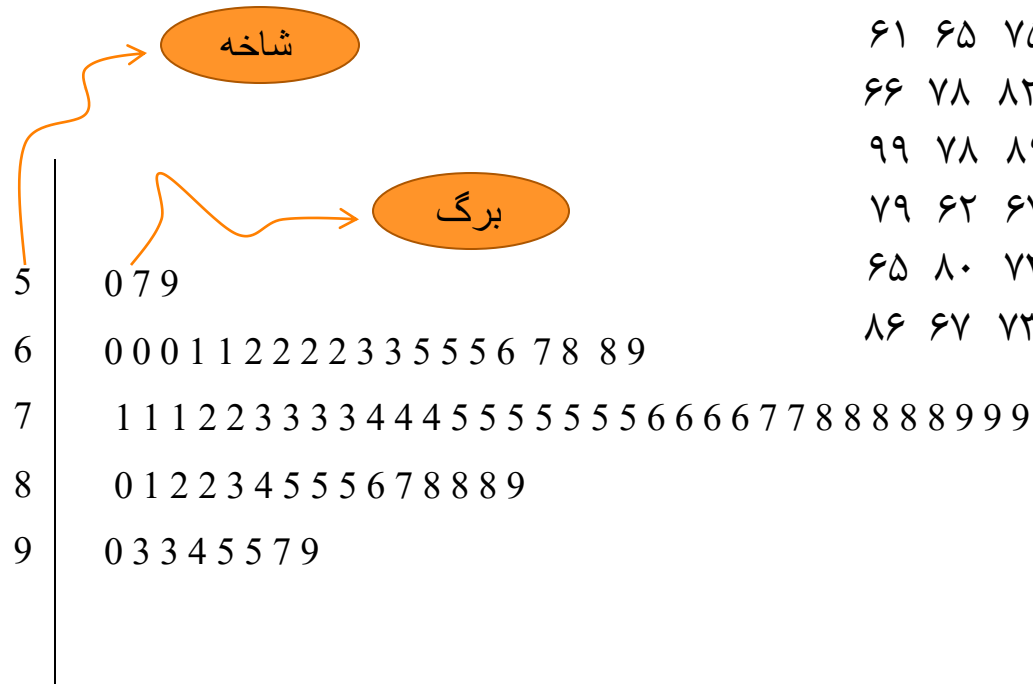
نمودار شاخه و برگ



نمودارهای آماری:

نمرات ۸۰ دانشجو در امتحانات نهایی درس آمار:

۶۸	۸۴	۷۵	۸۲	۶۸	۹۰	۶۲	۸۸	۷۶	۹۳
۷۳	۷۹	۸۸	۷۳	۶۰	۹۳	۷۱	۵۹	۸۵	۷۵
۶۱	۶۵	۷۵	۸۷	۷۴	۶۲	۹۵	۷۸	۶۳	۷۲
۶۶	۷۸	۸۲	۷۵	۹۴	۷۷	۶۹	۷۴	۶۸	۶۰
۹۹	۷۸	۸۹	۶۱	۷۵	۹۵	۶۰	۷۹	۸۳	۷۱
۷۹	۶۲	۶۷	۹۷	۷۸	۸۵	۷۶	۶۵	۷۱	۷۵
۶۵	۸۰	۷۳	۵۷	۸۸	۷۸	۶۲	۷۶	۵۰	۷۴
۸۶	۶۷	۷۳	۸۱	۷۲	۶۳	۷۶	۷۵	۸۵	۷۷



شاخص‌های مرکزی:

میانگین:

مهم‌ترین و مفیدترین شاخص مرکزی
متوسط (معدل) مقادیر داده‌ها
جمع کل داده‌ها تقسیم بر تعداد

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n f_i x_i$$

ویژگی‌ها:

استفاده از همه داده‌ها در محاسبه
اگر همه داده‌ها با یک عدد خاص جمع، تفریق، ضرب یا تقسیم شوند میانگین به همان نسبت تغییر می‌کند.

مجموع انحراف داده‌ها از میانگین صفر است.

موارد استفاده:

برای داده‌های کمی

شاخص‌های مرکزی:

میانہ:

مقداری که نصف داده‌ها از آن کوچکتر باشند (مقدار میانی یا وسطی)

اگر داده‌ها را از کوچک به بزرگ مرتب نماییم، عدد m را میانہ این داده‌ها می‌نامیم اگر نصف داده‌ها در سمت چپ و نصف داده در سمت راست این عدد قرار گیرد.

موارد استفاده:

داده‌های کمی یا کیفی ترتیبی

برتری نسبت به میانگین:

حساس نبودن به مقادیر بسیار بزرگ یا کوچک و متفاوت با سایر داده‌ها

قابلیت استفاده برای متغیرهای ترتیبی

شاخص‌های مرکزی:

نما:

داده‌ای که فراوانی آن نسبت به دیگر داده‌ها بیشتر باشد، (M)

ویژگی:

سادگی محاسبه

موارد استفاده:

برای همه‌ی انواع داده‌ها

شاخص‌های مرکزی:

چارک‌ها

چارک‌های اول، دوم و سوم (Q_1 ، Q_2 و Q_3) به ترتیب داده‌هایی هستند که $۰/۲۵$ ، $۰/۵۰$ و $۰/۷۵$ داده‌ها از آن کوچک‌تر هستند.

دهک‌ها

دهک‌های اول تا نهم (D_1 ، D_2 ، ...، D_9) به ترتیب داده‌هایی هستند که $۰/۱$ ، $۰/۲$ ، ...، $۰/۹$ داده‌ها از آن کوچک‌تر هستند.

دهک‌ها

صدک‌های اول تا نود و نهم (P_1 ، P_2 ، ...، P_{99}) به ترتیب داده‌هایی هستند که $۰/۰۱$ ، $۰/۰۲$ ، ...، $۰/۹۹$ داده‌ها از آن کوچک‌تر هستند.

شاخص‌های پراکندگی:

اهمیت با ذکر مثال:

دو مجموعه داده:

۱۱، ۱۲، ۱۳، ۱۴، ۱۵

۶، ۷، ۱۳، ۱۹، ۲۰

میانگین و میانه برابر

اما تفاوت‌های ساختاری زیاد

لازم است علاوه بر اندازه‌گیری شاخص‌های مرکزی، در خصوص تغییرات، تفاوت‌ها و اختلاف‌های بین داده‌ها نیز کسب اطلاع نمود.

بدین منظور از شاخص‌های پراکندگی استفاده می‌شود.

شاخص‌های پراکندگی

روشی ساده برای اندازه‌گیری پراکندگی داده‌ها

مطلوب نبودن در بسیاری از موارد

تنها استفاده از دو داده

$$R = y_{(n)} - y_{(1)}$$

دامنه

متوسط توان دوم اختلاف داده‌ها از میانگین

$$s^2 = \frac{1}{n} \sum_{i=1}^k f_i (x_i - \bar{x})^2$$

واحد اندازه‌گیری: توان دوم واحد اندازه‌گیری داده‌ها

واریانس

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^k f_i (x_i - \bar{x})^2}$$

جذر مثبت واریانس

انحراف معیار

واحد اندازه‌گیری مشابه داده‌ها

$$v = \frac{s}{\bar{x}}$$

ضریب تغییرات

شاخص پراکندگی بدون واحد

مقایسه پراکندگی متغیرهای با واحد اندازه‌گیری متفاوت

تعریف:

نمونه گیری فرایند انتخاب کردن تعداد کافی از میان اعضای جامعه آماری است، به طوری که با مطالعه گروه نمونه و فهمیدن خصوصیات یا ویژگیهای آزمودنیهای گروه نمونه قادر خواهیم بود این خصوصیات یا ویژگیها را به اعضای جامعه آماری تعمیم دهیم.

دلایل نمونه گیری:

در بررسی های پژوهشی که شامل چند صد و حتی چند هزار عضو جامعه آماری می شود عملا غیر ممکن است که اطلاعات را از هر عضو جمع آوری کنیم. حتی اگر امکان پذیر هم باشد به لحاظ زمان، هزینه و سایر مسایل منابع انسانی مقدور نیست.

علاوه بر اینها مطالعه یک گروه نمونه به جای کل جامعه آماری گاهی ممکن است منجر به نتایج معتبرتری شود به خاطر اینکه خستگی کمتری وجود خواهد داشت و از این رو خطاهای کمتری در جمع آوری اطلاعات پدید می آورد، مخصوصا موقعی که اعضای جامعه آماری بسیار وسیع باشد.

معرف بودن نمونه:

ما باید بتوانیم گروه نمونه را به صورتی انتخاب کنیم که معرف جامعه آماری باشد و ویژگیهای آن رادر برداشته باشد. به ندرت گروه نمونه جانشین کاملی برای جامعه آماری که از آن بیرون آمده است خواهد بود؛ ولی اگر گروه نمونه را به طور علمی برگزینیم می توانیم به طور منطقی مطمئن باشیم که آماره‌های گروه نمونه خیلی نزدیک به پارامترهای جامعه آماری خواهد بود.

نرمال بودن توزیع نمونه:

بسیاری از صفت ها یا ویژگیها در جامعه آماری عموماً به طور طبیعی توزیع می شود یعنی صفت هایی مانند قد و وزن چنان هستند که بیشتر مردم اطراف میانگین قرار دارند. اگر قرار باشد ویژگیهای جامعه آماری را با دقت معقول از ویژگی های موجود در گروه نمونه برآورد کنیم گروه نمونه باید چنان انتخاب شده باشد که توزیع ویژگیهای مورد نظر ما در آن همان نوع توزیع نرمال را که در جامعه وجود دارد دنبال کند. برای رسیدن به این هدف و داشتن یک گروه نمونه معرف باید اولاً حجم نمونه به حد کافی بزرگ بگیریم و ثانیاً طرح نمونه برداری مناسبی را انتخاب کنیم.

انواع نمونه گیری:

دو نوع اصلی طرح نمونه گیری وجود دارد:

۱. نمونه برداری احتمالی

۲. نمونه برداری غیر احتمالی

نمونه گیری احتمالی:

در نمونه گیری احتمالی اعضای جامعه شانس یا احتمال شناخته شده ای دارند که به عنوان آزمودنی گروه نمونه انتخاب شوند.

طرح های نمونه گیری احتمالی موقعی به کار می رود که معرف بودن گروه نمونه برای اهداف تعمیم پذیری دارای اهمیت باشد.

✓ نمونه گیری احتمالی شامل دو دسته نامحدود و محدود می باشد.

نمونه گیری احتمالی نا محدود (نمونه برداری تصادفی ساده):

در این نوع نمونه گیری اعضای جامعه آماری یک شانس معین و برابر برای انتخاب شدن به عنوان آزمودنی دارند. این نوع نمونه گیری کمترین سوگیری و بیشترین تعمیم پذیری را دارا می باشد. اما این روش پر زحمت و پرهزینه است و گاهی نمی توانیم فهرست کاملاً جدیدی از جامعه آماری بدست آوریم.

نمونه گیری احتمالی محدود یا پیچیده:

مرسوم‌ترین طرح‌های نمونه گیری احتمالی پیچیده که به عنوان جایگزین‌هایی برای طرح
پرزحمت و پر هزینه احتمالی ساده به شمار می‌روند عبارتند از:

الف) نمونه گیری سیستماتیک

ب) نمونه گیری تصادفی طبقه‌ای

ج) نمونه گیری خوشه‌ای

د) نمونه گیری خوشه‌ای مرحله‌ای

الف) طرح نمونه گیری نظام دار(سیستماتیک):

طرح نمونه گیری منظم شامل بیرون آوردن هر n امین عضو جامعه آماری است, به گونه‌ای که این عضو به طور تصادفی بین 1 و n انتخاب می شود. اگر ما گروه نمونه ای از 35 خانوار از کل جامعه آماری شامل 260 خانه در یک محل خاص را بخواهیم, می توانیم با انتخاب شماره تصادفی مثلا 7 هر هفتمین خانه را با شمارش 1 تا 7 نمونه برداری کنیم یعنی خانه های شماره 7 و 14 و 21 و 28 ...

ب) طرح نمونه گیری تصادفی طبقه ای:

نمونه گیری تصادفی طبقه‌ای مستلزم طبقه‌بندی جامعه آماری و سپس انتخاب تصادفی آزمودنی‌ها از هر طبقه است. در نمونه برداری تصادفی طبقه ای جامعه آماری ابتدا به گروه‌های ناسازگاری که در بافت پژوهش مرتبط، متناسب و معنادار هستند تقسیم می‌شود. و سپس از هر طبقه نمونه برداری صورت می‌گیرد. پیدا کردن احتمالات در پارامترهای گروه‌های فرعی در داخل جامعه آماری بدون روش نمونه گیری تصادفی طبقه ای امکان پذیر نیست. وقتی جامعه آماری با روش معناداری طبقه‌بندی شد، نمونه گیری منظم انتخاب شود.

پ) نمونه گیری خوشه ای :

این طرح شامل تقسیم بندی جامعه آماری به خوشه‌های مناسب میشود، در حالی که بطور تصادفی تعداد لازم از خوشه‌ها را به عنوان آزمودنی‌های گروه نمونه انتخاب کرده و پس از آن هم اعضای جامعه آماری در هر یک از این خوشه‌ها مورد مطالعه قرار می‌گیرند.

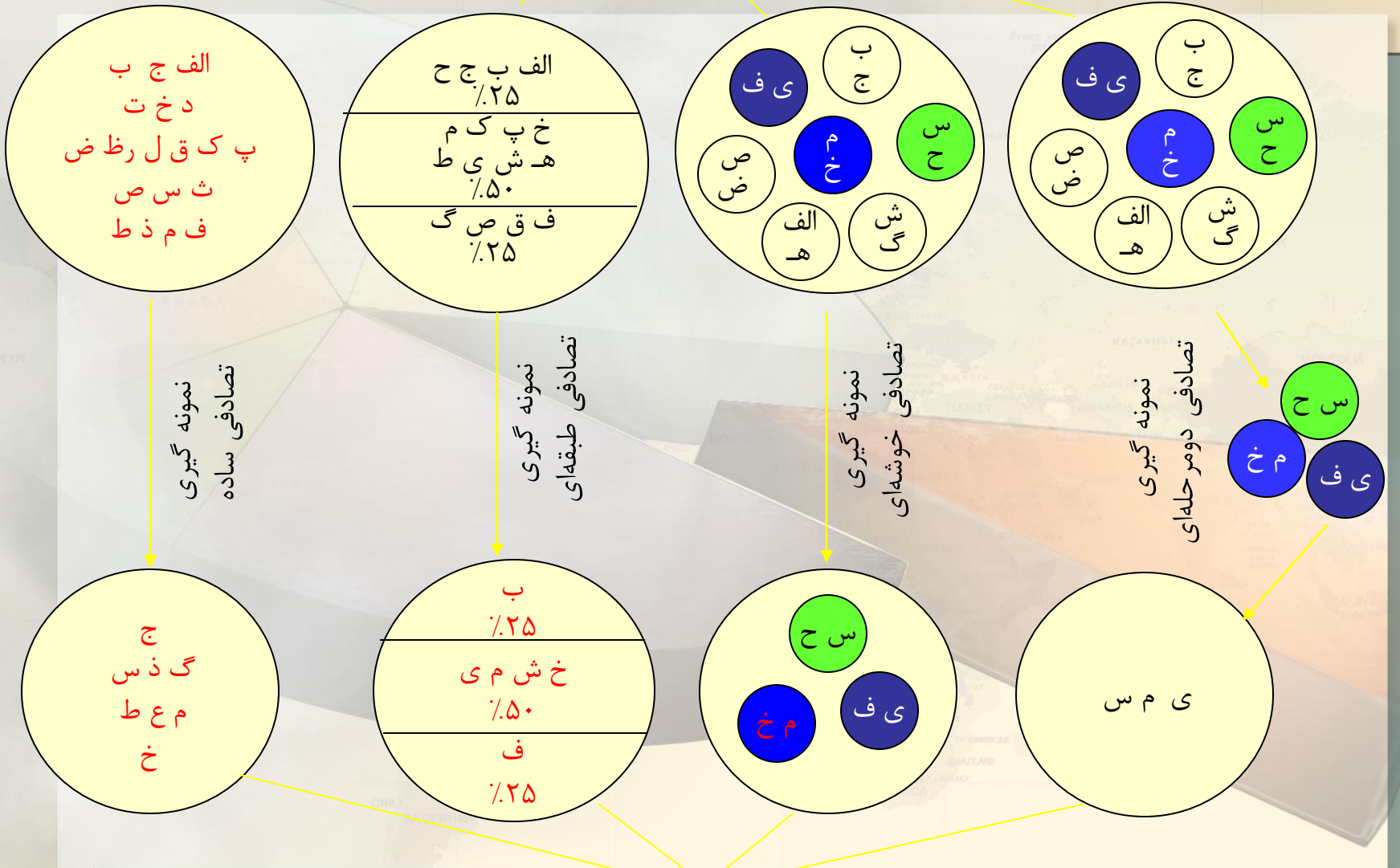
گروه‌هایی از اعضای جامعه آماری به طوری که عدم تجانس در میان اعضای هر گروه وجود داشته باشد برای مطالعه انتخاب میشوند. نمونه گیری خوشه‌ای عدم تجانس بیشتری را در داخل گروه‌ها و تجانس بیشتری را در بین گروه‌ها ارائه می‌کند که عکس آن چیزی است که ما در نمونه گیری تصادفی طبقه‌ای می‌یابیم زیرا در آن تجانس در هر گروه و عدم تجانس در میان گروه‌ها وجود دارد.

نمونه گیری خوشه‌ای چند مرحله‌ای Multiple cluster sampling:

بواسطه گستردگی بیش از حد جامعه محقق ناگزیر می‌گردد نمونه را طی دو یا چند مرحله انتخاب کند .

- جامعه موردنظر را دقیقاً تعریف کنید .
- واحدها یا خوشه‌های نمونه برداری را تعریف کنید .
- تعدادی از خوشه‌ها یا واحدها را به صورت تصادفی انتخاب کنید .
- از میان خوشه‌های انتخاب شده تعداد افراد مورد نظر را به روش تصادفی انتخاب کنید .

جامعه



نمونه گیری غیر احتمالی:

اعضای جامعه آماری هیچ احتمالی برای انتخاب شدن در گروه نمونه به عنوان آزمودنی را ندارند. بدین معنی که یافته های مطالعه گروه نمونه را نمی توان به اطمینان به جامعه آماری تعمیم داد. و زمانی مورد استفاده قرار می گیرد که پژوهشگر به روش سریع و ارزان بیشتر علاقه من باشد تا تعمیم پذیری یافته ها.


✓ نمونه گیری غیر احتمالی دو گروه عمده نمونه برداری در دسترس و نمونه برداری هدف دار را شامل می شود .

الف) نمونه برداری در دسترس:

در این طرح، در دسترس ترین اعضا به عنوان آزمودنی انتخاب می شوند این طرح به هیچ وجه قابل تعمیم نمی باشد و برای پژوهش علمی مناسب نیست و ممکن است در مواقعی برای کسب اطلاعات سریع و استفاده از نتایج آن به کار رود.

ب) نمونه گیری هدف دار:

ممکن است گاهی ضروری باشد به جای کسب اطلاعات از کسانی که در دسترس هستند، اطلاعات را از افراد خاص بدست آوریم یعنی افرادی که قادر خواهند بود اطلاعات مطلوب را ارائه دهند به این دلیل که آنها تنها کسانی هستند که می توانند اطلاعات لازم را بدهند یا افرادی هستند که با معیار خاص که پژوهشگر در نظر دارد وفق دهد. چنین روش نمونه گیری را هدفمند می نامند.



انواع نمونه گیری هدفدار:

• نمونه برداری قضاوتی

• نمونه برداری سهمیه‌ای

۱) نمونه برداری قضاوتی:

این طرح زمانی به کار می رود که گرد آوری "داده های تخصصی و دقیق" در نمونه موضوع پژوهش جنبه حیاتی دارد و هیچ یک از طرح های نمونه برداری احتمالی امکان گرد آوری چنین داده هایی را بدست نمی دهد. نمونه گیری قضاوتی مستلزم انتخاب آزمودنی هایی است که بهترین شرایط را برای ارائه اطلاعات مورد نیاز دارند.

۲) نمونه برداری سهمیه ای:

نمونه گیری سهمیه ای شکلی از نمونه گیری طبقه ای متناسب با حجم است که در آن از افراد گروه های مختلف و در دسترس را بر می گزینند. یعنی آزمودنی ها منحصر از اعضای مورد نظر و در دسترس بر طبق تعداد یا سهمیه ای که پیشتر تعیین شده انتخاب می شوند.

برآورد حجم نمونه:

۱. حجم نمونه برای برآورد یک نسبت در جمعیت:

$$n = \frac{z_{1-\alpha/2}^2 \times p(1-p)}{d^2}$$

$$d = 0.1p$$

مثال: محقق می‌خواهد برآورد نماید که احتمال یک خطر ۱۰ ساله بیماری قلبی عروقی در بین مردان ۴۰ تا ۷۹ ساله در شمال ایران چقدر است. در مطالعه‌ای در سال ۲۰۰۹ که نتایج آن در سال ۲۰۱۵ منتشر شد این شیوع برابر ۴۸/۹ درصد برآورد گردید. این محقق، به چند نمونه نیاز خواهد داشت، اگر خطای نوع اول را برابر ۵ درصد در نظر بگیرد؟

بر آورد حجم نمونه:

۲. حجم نمونه برای برآورد یک میانگین در جمعیت:

$$n = \frac{z_{1-\alpha/2}^2 \times \sigma^2}{\delta^2}$$

مثال: محقق می‌خواهد میانگین فشار خون سیستولیک زنان ۱۸ سال و بالاتر را در شمال ایران تخمین بزند. براساس مطالعه ای که در گذشته بر روی همین جمعیت صورت گرفت برآورد فشار خون برابر $17/60 \pm 115/42$ بدست آمد. برای هدایت این مطالعه محقق به چند نمونه نیاز خواهد داشت، اگر دقت را در این مطالعه ۲ میلی متر جیوه در نظر گرفته باشد (البته با خطای یک درصد)؟

برآورد حجم نمونه:

۳. برآورد حجم نمونه برای مقایسه میانگینهای دو جامعه با واریانسهای برابر:

$$n_A = \frac{(\varphi + 1)(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\varphi \delta^2}$$

$$n_B = \varphi n_A$$

۴. برآورد حجم نمونه برای مقایسه میانگینهای دو جامعه با واریانسهای نابرابر:

$$n_{A(\text{Variances Unequal})} = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \left(\sigma_A^2 + \frac{\sigma_B^2}{\varphi} \right)}{\delta^2}$$

$$n_{B(\text{Variances Unequal})} = \varphi \times n_{A(\text{Variance Unequal})}$$

مثال: فرض کنید محققى بخواهد میانگین فشار خون را در دو جمعیت مقایسه کند. او این فرضیه را در ذهن خود دارد که میانگین فشار خون در این دو جامعه با هم برابر نیست. اگر واریانسها باهم برابر باشند یا نباشند چه مقدار حجم نمونه مورد نیاز است در صورتی که انحراف معیارهای دو جامعه براساس برآوردهای قبلی برابر ۱۵ و ۲۰ میلیمتر جیوه باشد، خطای آلفا ۵ درصد، توان مطالعه ۹۰ درصد و حداقل تفاوت با اهمیت از نظر بالینی برابر ۱۰ میلیمتر جیوه اختیار شود. این محقق نسبت تخصیص را برابر ۱ میگیرد.

بر آورد حجم نمونه:

۵. بر آورد حجم نمونه برای مقایسه یک میانگین با یک عدد مرجع:

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\varepsilon^2}$$

$$\varepsilon = |\mu_1 - \mu_0|$$

مثال: فرض کنید محقق می‌خواهد تعیین کند آیا تفاوت معنی داری بین میانگین قند خون جامعه ای با یک عدد ثابت که آن را ۱۰۰ میلی گرم در دسی لیتر در نظر گرفت، وجود دارد یا خیر؟ فرض کنید او ۱۰ میلی‌گرم در دسی لیتر تفاوت در قند خون ناشتا را از لحاظ بالینی با اهمیت تلقی کرده و می‌خواهد بداند که آیا قند خون اندازه گیری شده بیش از ۱۰ میلی گرم با عدد ۱۰۰ تفاوت دارد یا خیر؟ محقق خطای آلفا را ۵ درصد و خطای بتا را ۲۰ درصد برای مطالعه در نظر می‌گیرد.

بر آورد حجم نمونه:

۶. بر آورد حجم نمونه برای مقایسه میانگینهای دو جامعه مستقل که نتایج در دو زمان متفاوت بدست میآیند:

$$n = \frac{2\sigma_d^2 \times (z_{1-\alpha/2} + z_{1-\beta})^2}{\delta^2}$$

$$\sigma_d^2 = \sigma_1^2 + \sigma_2^2 - 2\rho \times \sigma_1 \times \sigma_2 \quad \delta = |\mu_1 - \mu_2|$$

مثال: محققى قصد دارد تغییرات میانگین فشارخونهای سیستولیک را در بین گروه درمانی و کنترل مقایسه نماید. براساس مطالعات قبلی انحراف معیارهای مقادیر پایه و پیگیری (با فاصله یکسال) به ترتیب برابر ۱۵ و ۱۲ میلی متر جیوه وضریب همبستگی بین دو اندازه گیری برابر ۰/۷ بدست آمد. همچنین کاهش میانگین برای گروه درمان در عرض ۱ سال ۱۰ میلیمتر جیوه و گروه کنترل ۳ میلیمتر جیوه بدست آمد. اگر این محقق بخواهد چنین مطالعه ای را با توان ۸۰ درصد و خطای نوع اول ۵ درصد هدایت نماید به چه حجم نمونه ای نیاز خواهد داشت؟