



درس تحلیل آماری

مدرس: دکتر سمیه عباسی

عضو هیات علمی دانشگاه آزاد اسلامی اصفهان (خوراسگان)

دانشگاه آزاد اسلامی
واحد اصفهان خوراسگان



واحد اصفهان خوراگان

مدرس: دکتر سمیه عباسی

عنوان درس: تحلیل آماری-فصل اول: آمار توصیفی

تحلیل آماری

- ۱- میان ترم، ۴ نمره
- ۲- پایان ترم، ۱۴ نمره
- ۳- حل تمرین و مشارکت در کلاس، ۲ نمره

منبع درسی
آمار و کاربرد آن در مدیریت، دکتر
عادل آذر و دکتر منصور مومنی
جلد دوم

سر فصل

- مفاهیم مقدماتی
- توزیع نرمال
- آزمون فرض
- تعیین حجم نمونه
- رگرسیون و همبستگی
- آنالیز واریانس
- آزمون های ناپارامتری
- آشنایی با spss

فصل اول: آمار توصیفی

تعریف علم آمار

مجموعه ای از روش‌ها و تکنیک‌های علمی جهت:

پردازش داده‌ها

توصیف داده‌ها

جمع آوری داده‌ها

و همچنین

اطلاعات مورد نیاز در جامعه

برای رسیدن به

استفاده از
داده‌های نمونه

پس بر این اساس آمار به دو
دسته تقسیم می شود

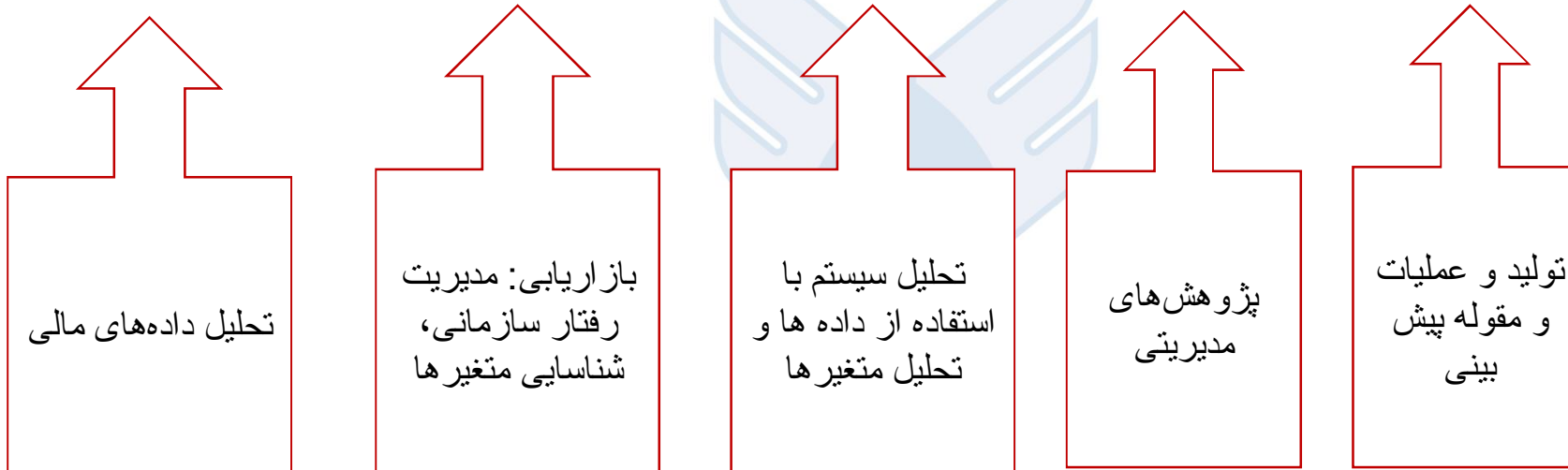
آمار استنباطی:

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

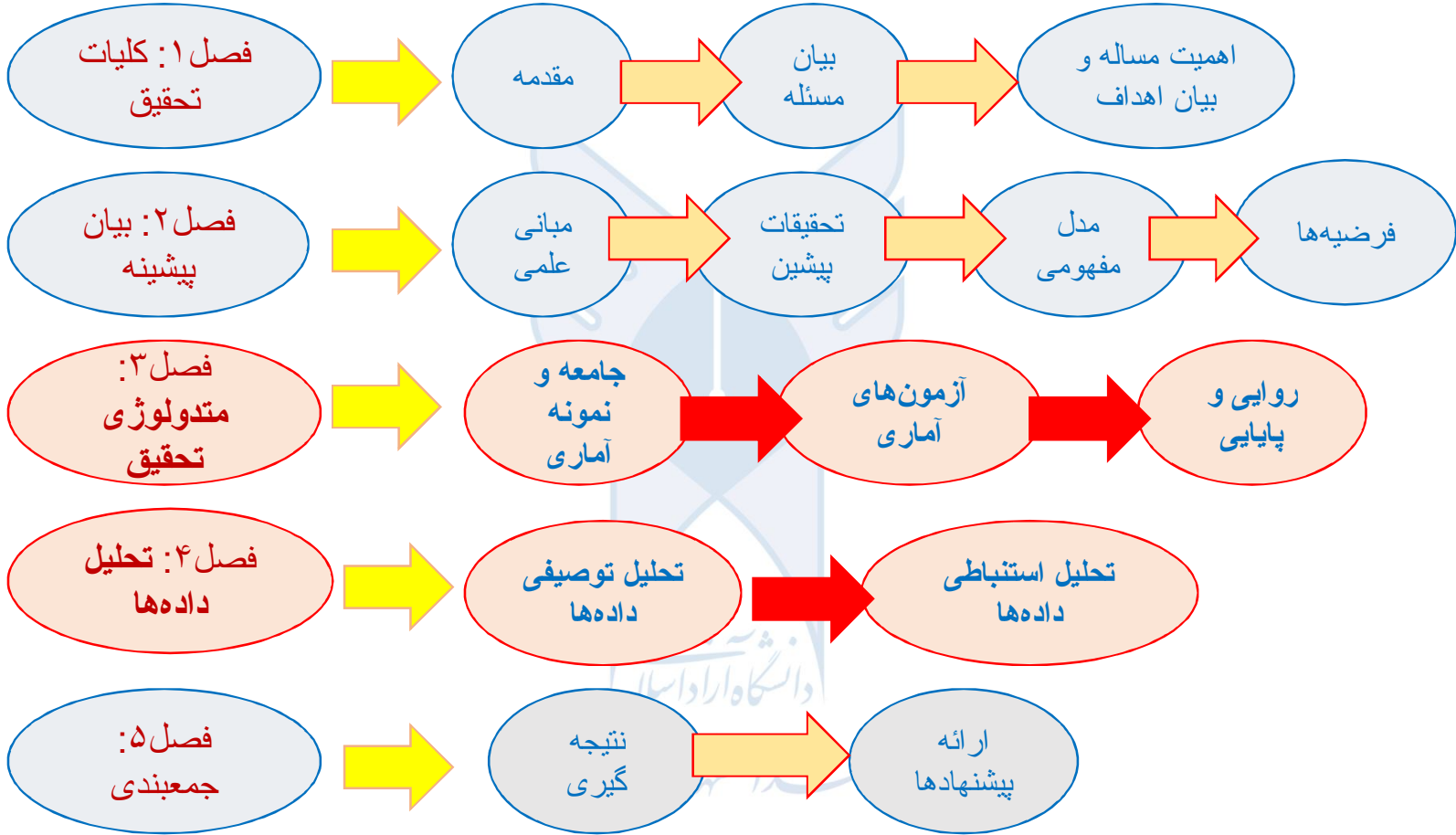
آمار توصیفی:

خلاصه کردن و توصیف
خصوصیات مهم جامعه یا
نمونه مورد بررسی در
قالب جداول و نمودار به
کمک شاخص های مرکزی
و پراکنندگی

لزوم کاربرد آمار در مدیریت



مرحل اساسی انجام پژوهش علمی و نیاز به
آمار در فصول ۳ و ۴



جامعه آماری

جامعه آماری که به آن جمعیت آماری نیز گفته اند به مجموعه افراد یا عناصر (بستگی به موضوع تحقیق دارد) که دارای ویژگی یا ویژگی های مشترکی هستند و تحقیق در مورد آنها انجام می شود گفته می شود. در متون آماری حجم جامعه با N نمایش داده می شود. مانند کارمندان فولاد مبارکه

سوال

اگر قصد بررسی نظر رای دهندگان در مورد کاندیدای نمایندگی یک شهر داشته باشیم جامعه آماری ما کدام است؟

سرشماری

جمع آوری اطلاعات از همه اعضای جامعه را سرشماری می گوئیم. بنابراین سرشماری بررسی است که تمام اعضا و واحد های جامعه را در بر می گیرد.

سوال

آیا در همه تحقیقات می توان کلیه جامعه آماری را مورد آزمایش و بررسی قرار داد؟

مهمترین محدودیت‌ها

- کمبود وقت
- هزینه بر بودن
- غیر ممکن و غیر معقول بودن

نمونه آماری
Sample

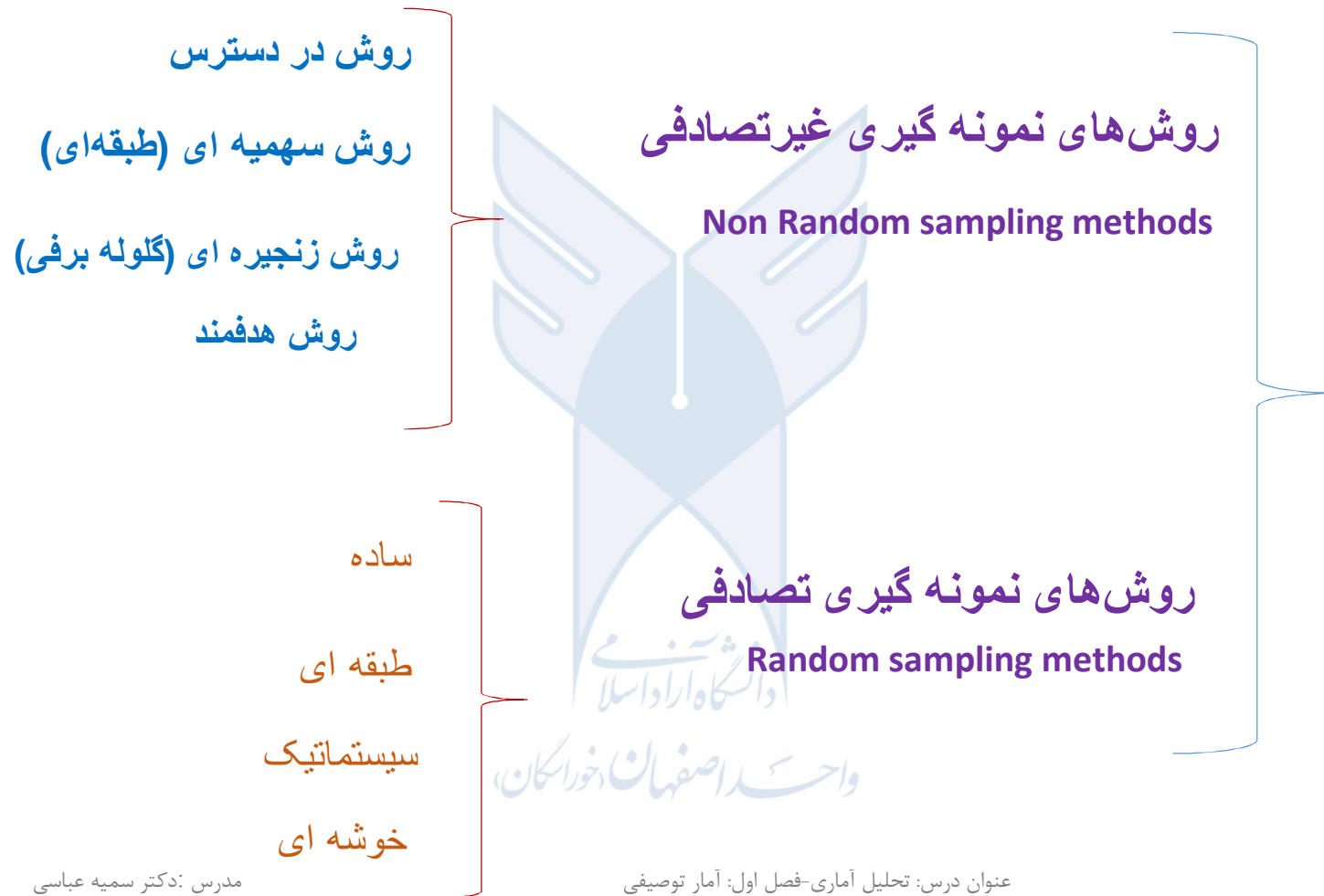
بخشی (زیرمجموعه ای) از جامعه تحت بررسی (جامعه آماری) است که با روشی که از پیش تعیین شده است (روش نمونه گیری) انتخاب می شود؛ به قسمی که می توان از این بخش (نمونه) استنباطهایی درباره کل جامعه به دست آورد. حجم نمونه با n نشان داده می شود. مانند کارمندان بخش حسابداری فولاد مبارکه

هزینه انتخاب نمونه کم باشد

نمونه باید نماینده جامعه باشد

دقت نمونه گیری بالا باشد

روشهای نمونه گیری Sampling methods



روش های نمونه گیری غیر تصادفی

در برخی پژوهش ها بسیار اتفاق می افتد که پژوهشگر نمی تواند نمونه ای انتخاب کند که معرف جامعه باشد. موارد استفاده آن در مطالعات مقدماتی یا در پژوهش هایی است که برای بررسی مشکلات احتمالی یک پژوهش خاص در محیطی محدود می باشد.

در این نوع نمونه گیری، پژوهشگر از نمونه های در دسترس استفاده می کنند. مثل استفاده پزشک از بیماران مراجعه کننده به مطب شخصی خود برای مقایسه اثر دو دارو در درمان یک بیماری یا استفاده یک حسابرس از فاکتورهای موجود در کنترل نامحسوس حسابداری یک ارگان

روش در دسترس

در این نوع نمونه گیری، پژوهشگر از هر یک از گروه ها به نسبت وزنی که در جامعه دارند، نمونه هایی را از میان افراد در دسترس انتخاب می کند. مانند بررسی میزان درآمد کادر بیمارستانی که تعداد جمعیت پرستاران سه برابر جمعیت پزشکان است و براساس این روش نسبت فوق در انتخاب افراد نمونه رعایت می شود.

روش سهمیه ای

روشی است که در آن شرکت‌کنندگان در پژوهش، شرکت‌کنندگان اضافی و دیگری را برای مطالعه پیشنهاد می‌نمایند. به طور مثال در پژوهش‌های کیفی با روش مصاحبه، از مصاحبه‌شوندگان سؤال می‌شود که آیا فرد دیگری را برای انجام مصاحبه پیشنهاد می‌نمایید که بتواند اطلاعات مناسبی در این خصوص ارائه دهد.

روش
زنجیره‌ای

این روش نوعی گلچین کردن افراد از میان یک جمعیت است. در این روش پژوهشگر بر این باور است که از بصیرت و آگاهی خود می‌تواند برای شناسایی و انتخاب افراد مناسب برای مصاحبه یا مشاهده استفاده نماید. به عبارت دیگر در این روش نمونه‌ها بر اساس قضاوت پژوهشگر انتخاب می‌شود و به همین دلیل به آن روش قضاوتی نیز اطلاق می‌شود.

روش
هدفمند

روش‌های نمونه‌گیری تصادفی

روش تصادفی ساده

در این روش تمام اعضای جامعه دارای شانس مساوی یا حداقل شانس معینی برای قرار گرفتن در نمونه خواهند بود.
در عمل همه نمونه‌گیری‌ها بدون جایگزینی انجام می‌شود. در نمونه‌گیری تصادفی ساده شانس انتخاب هر نمونه n/N است. استفاده از قرعه‌کشی و جدول اعداد تصادفی در این روش بسیار شایع است.

- مزایای نمونه‌گیری تصادفی ساده
- سادگی روش
- خطای نمونه‌گیری به سادگی قابل اندازه‌گیری است
- برای نمونه‌گیری از جوامع کوچک روشی مناسب و عملی است.
- سوگیری انتخاب در این روش وجود ندارد.



معایب نمونه‌گیری تصادفی ساده

- نیاز به لیست کامل و روز آمدی از جامعه دارد.
- نمونه ممکن است بسیار پراکنده باشند و دسترسی به همه نمونه‌ها امکان‌پذیر نباشد.
- هزینه‌بر است.
- به خوبی معرف جامعه دارای سطوح مختلف متغیر نیست.

روش سیستماتیک یا منظم

ایده‌ی اساسی نمونه‌گیری سیستماتیک انتخاب منظم تعدادی از افراد موجود در یک لیست می‌باشد.

فرض کنید قرار است نمونه‌ای از اسم از فهرستی طولانی انتخاب شود. یک روش ساده برای این کار آن است که فاصله‌ای مناسب برگزیده و اسامی را در فواصل مساوی در طول فهرست انتخاب کنیم. اگر نقطه‌ی شروع برای این روند منظم انتخاب، تصادفی باشد نتیجه یک نمونه‌گیری سیستماتیک است.

به این ترتیب نمونه‌ای که به وسیله‌ی انتخاب یک عضو از بین اولین عضو یک فهرست، سپس انتخاب هر k مین عضو بعد از آن حاصل شود، یک نمونه تصادفی سیستماتیک ۱ در k با شروع تصادفی نامیده می‌شود.

این روش برای آن دسته از جوامع آماری که کد از پیش تعیین شده و مرتبی دارند (همانند شماره کارمندی، دانشجویی و پلاک منازل) کاربرد فراوان دارد. با مشخص شدن اولین عضو نمونه، سایر اعضای نمونه در این روش معین می‌شوند.

در نمونه‌گیری با روش منظم تنها یک عدد تصادفی در کل پروسه نمونه‌گیری لازم است که این خاصیت از یک سو یکی از محاسن روش تلقی و از سوی دیگر موجب از دست رفتن شانس انتخاب برای سایر اعضای جامعه می‌شود.

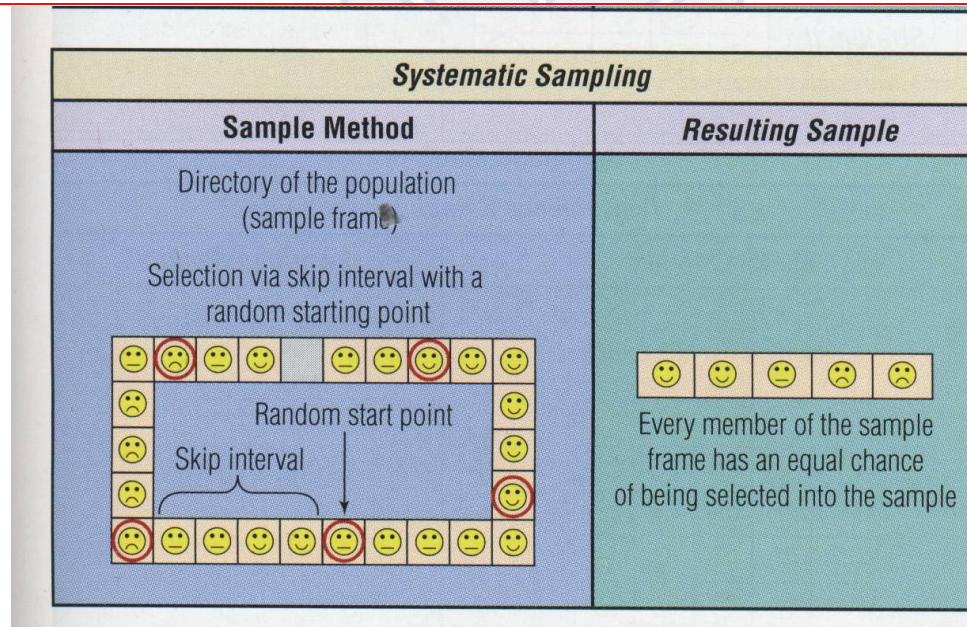
روش انجام نمونه‌گیری

- ۱- حجم نمونه تعیین می‌شود.
- ۲- فاصله نمونه‌گیری تعیین می‌شود.
$$\frac{N}{n} = \text{فاصله نمونه‌گیری}$$
- ۳- انتخاب اولین نمونه بر اساس روش تصادفی بین ۱ تا عدد فاصله نمونه‌گیری صورت می‌گیرد.
نمونه‌های بعدی با اضافه نمودن فاصله نمونه‌گیری به نمونه ماقبل بدست می‌آیند.

مثال:
در یک بیمارستان ۱۰۰۰ تخت خوابی قرار است ۵۰ بیمار جهت بررسی علل ابتلا به کرونا
انتخاب شوند:

$$\text{فاصله نمونه گیری} = 1000 / 50 = 20$$

فرض کنید بطور تصادفی اولین نمونه عدد ۸ انتخاب شده است.
نمونه دوم = $20 + 8 = 28$
نمونه سوم = $28 + 20 = 48$





با توجه به معایب نمونه‌گیری تصادفی ساده و منظم، در صورت امکان تقسیم جامعه مورد پژوهش به طبقه‌ها یا گروه‌ها می‌توان نمونه‌گیری را در هر طبقه انجام داد

در نمونه‌گیری تصادفی ساده و منظم نمی‌توان اطمینان داشت که در نمونه مورد نظر، نسبت کسانی که دارای ویژگی‌های معین هستند برابر همان نسبت در جامعه باشند. برای حل این مشکل، ممکن است با استفاده از اطلاعات کمکی بتوان جامعه را به طبقه‌ها یا گروه‌هایی تقسیم کرد در آن صورت می‌توان نمونه‌گیری را در هر طبقه انجام داد البته زمانی می‌توانیم از این نوع نمونه‌گیری استفاده کنیم که نسبت هر طبقه در جامعه مشخص باشد. اگر بتوان جامعه مورد مطالعه را بر اساس صفت مشخصی طوری به گروه‌های جدا از هم (بدون همپوشانی) تقسیم کرد که تعداد افراد هر گروه معلوم باشد می‌توان نمونه‌گیری با طبقه-بندی را با انجام نمونه‌گیری تصادفی ساده در هر طبقه و ادغام نتیجه در هم انجام داد. در عین حال یکی از دلایل دیگر برای انتخاب این نوع نمونه‌گیری آن است که با استفاده از این نوع نمونه‌گیری می‌توان پارامترهای مربوط را برای هر طبقه برآورد کرد.

روش انجام

- ۱- جمعیت تحت مطالعه دارای N عضو است.
 - ۲- جمعیت از لحاظ داشتن صفات خاصی به H طبقه متفاوت تقسیم می‌شوند.
 - ۳- هر عضو فقط می‌تواند به یکی از طبقات تعلق داشته باشد. (بدون همپوشانی)
 - ۴- تعداد مشاهدات در هر طبقه مشخص است و با N_h مشخص می‌شود.
 - ۵- تعداد کل جمعیت برابر است با مجموع تعداد جمعیت‌ها در هر طبقه.
 - ۶- محقق با استفاده از به کارگیری یکی از روش‌های نمونه‌گیری مثل نمونه‌گیری تصادفی ساده، از هر طبقه تعداد نمونه مورد نیاز را انتخاب می‌کند.
- بعد از اینکه طبقات را مشخص کردیم برای انتخاب نمونه از هر طبقه ۲ راه وجود دارد:
- ۱- انتخاب تعداد مساوی از هر طبقه
 - ۲- انتخاب نمونه از هر طبقه با در نظر گرفتن نسبت آن طبقه در جمعیت

واحد اصفهان خوراگان

راه اول و ساده‌تر این است که تعداد مساوی از هر طبقه انتخاب کنیم $n_k = n/L$ که L برابر است با تعداد طبقات و n_h برابر است با تعداد عناصر انتخاب شده از طبقه h ام و n هم کل تعداد نمونه است.

مثال: یک جمعیت ۱۰۰۰ نفری را از نظر مقطع تحصیلی به ۴ طبقه تقسیم کرده‌ایم که طبقات شامل ۳۰۰، ۲۰۰ و ۱۰۰ هستند و می‌خواهیم از بین آنها ۱۰۰ نفر را به عنوان نمونه انتخاب کنیم تعداد افراد هر نمونه برابر می‌شود با $n_k = \frac{100}{4}$ و به این ترتیب از هر طبقه ما ۲۵ نفر را انتخاب می‌کنیم.

$$\cdot \frac{25}{100}, \frac{25}{200}, \frac{25}{300}, \frac{25}{400}$$

راه دوم: روش مفید تر انتخاب افراد در طبقات اختصاص تناسبی proportional Allocation است. در این روش کسر

نمونه‌گیری $\frac{n_k}{N_k}$ مختص هر طبقه است و در مثال قبل برای طبقات برابر است با:

$$n_h = N_h \left(\frac{n}{N} \right)$$

واحد رضوان خراسان

$$\bullet P_1 = (400) \times \left(\frac{100}{1000}\right) = 40$$

$$\bullet P_2 = (300) \times \left(\frac{100}{1000}\right) = 30$$

$$\bullet P_3 = (200) \times \left(\frac{100}{1000}\right) = 20$$

$$\bullet P_4 = (100) \times \left(\frac{100}{1000}\right) = 10$$



یک نمونه ی خوشه ای نمونه ای احتمالی است که در آن هر واحد نمونه گیری مجموعه ای یا گروهی از اعضا است. دلایل مختلفی برای استفاده از نمونه گیری خوشه ای می تواند وجود داشته باشد. در صورتی که هزینه ی فراهم نمودن یک چارچوب که کلیه اعضای جامعه را فهرست می کند زیاد باشد و یا اگر هزینه فراهم آوردن مشاهدات با افزایش مسافت بین اعضا افزایش یابد، نمونه گیری خوشه ای می تواند کم هزینه تر از نمونه گیری تصادفی ساده یا طبقه ای باشد.

واحد رصفهان خوراگان

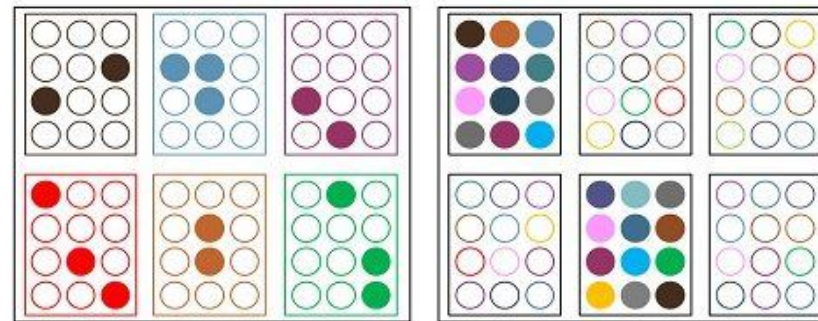
مقایسه نمونه گیری خوشه ای و طبقه ای

۱) در نمونه گیری تصادفی طبقه ای از هر طبقه تعدادی را به عنوان نمونه انتخاب می کنیم در صورتی که در نمونه گیری تصادفی خوشه ای، نمونه از تعدادی از خوشه ها انتخاب می شود.

۲) در نمونه گیری طبقه ای ، طبقات تا آنجایی که ممکن است باید از درون به هم شبیه باشند اما یک طبقه باید با طبقه ی دیگر از نظر مشخصه ای که اندازه گیری می شود متفاوت باشد. به عبارت دیگر دقت نمونه گیری تصادفی طبقه ای با همگنی درون طبقات و ناهمگنی بین طبقات بیشتر می شود.

۳) در نمونه گیری خوشه ای، خوشه ها تا حد امکان باید از درون متفاوت باشند و یک خوشه باید به منظور هدف مورد مطالعه، بسیار شبیه دیگر خوشه ها به نظر آید. به عبارت دیگر دقت نمونه گیری تصادفی خوشه ای با ناهمگنی درون خوشه ها و همگنی بین خوشه ها بیشتر می شود.

روش طبقه بندی



مدرس: دکتر سمیه عباسی

عنوان درس: تحلیل آماری-فصل اول: آمار توصیفی

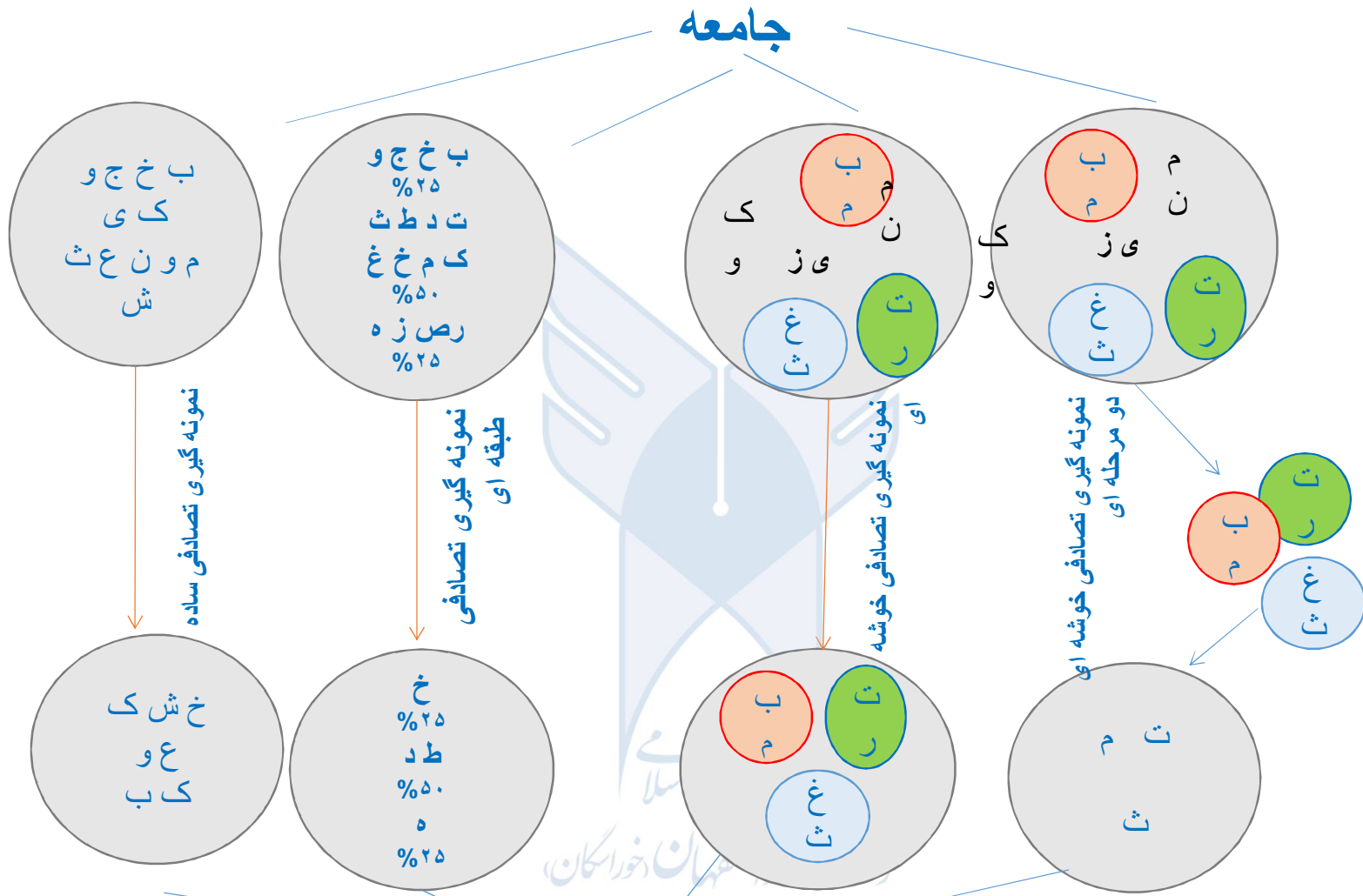
روش انجام

در نمونه گیری خوشه ای در مرحله ابتدایی باید از بین خوشه های تشکیل دهنده جامعه مورد بررسی تعدادی خوشه را انتخاب کنیم. نحوه انتخاب خوشه ها می تواند تصادفی ساده یا منظم باشد. پس از انتخاب خوشه های مورد نظر وارد مرحله بعدی این نوع نمونه گیری می شویم. در این مرحله دو انتخاب اساسی داریم. در انتخاب اول کلیه اعضای خوشه های منتخب را مورد بررسی و ارزیابی قرار می دهیم. به عبارت دیگر کلیه اعضای خوشه به عنوان حجم نمونه مطالعه در نظر گرفته می شوند به این شیوه نمونه گیری خوشه ای، نمونه گیری خوشه ای یک مرحله ای اطلاق می گردد. در انتخاب دوم از بین کلیه اعضای خوشه یا خوشه های انتخاب شده تعدادی از اعضا به شیوه تصادفی ساده یا منظم انتخاب می گردند به عبارت دیگر در این حالت در داخل خوشه ها نمونه گیری انجام می دهیم که به این شیوه، نمونه گیری خوشه ای دو مرحله ای می گویند. این روش نمونه گیری قابل بسط به تعداد مرحله های بیشتر نیز می باشد.

مثال

در مطالعه ارزیابی کیفیت آموزش مجازی در سامانه شاد در دانش آموزان ابتدایی شهر تهران یک شیوه نمونه گیری می تواند به صورت زیر باشد: ابتدا از بین ۲۲ منطقه شهر تعدادی از مناطق به صورت تصادفی ساده انتخاب می شوند (۵ منطقه)، لیستی از مدارس ابتدایی مناطق پنج گانه انتخاب شده تهیه می کنیم و سپس ۱۰ دبستان در هر منطقه به صورت تصادفی ساده انتخاب می گردد و در نهایت در هر دبستان انتخاب شده ۵۰ دانش آموز به صورت تصادفی منظم براساس لیست حروف الفبایی انتخاب می گردند که با کمی دقت در می یابیم این شیوه نمونه گیری، نمونه گیری خوشه ای سه مرحله ای می باشد.

ترسیمی از نمونه گیری تصادفی



نمونه

عنوان درس: تحلیل آماری-فصل اول: آمار توصیفی

مدرس: دکتر سمیه عباسی

نمودارهای آماری:

معمولا داده‌ها را با نمودارهاي مختلف نمایش مي‌دهند. عموما این نمودارها در ارتباط با داده‌هاي پیوسته به کار گرفته مي‌شود و منظور از نمایش آنها، **تجسم عيني اطلاعات نهفته در داده‌ها است**. در این بخش به معرفي چند نمودار معروف اکتفا مي‌کنیم:

- نمودار ستونی

== هیستوگرام

== چندبر فراوانی

معیارهای مرکزی

- با استفاده از جدول فراوانی و رسم نمودارها می‌توانیم داده‌ها را به نحو مطلوبی **تنظیم کرده و اطلاعات نهفته را تا حدودی مشخص کنیم**. با این حال برای آرایه یک گزارش مناسب، بهتر است آنها را در یک یا چند عدد مناسب نیز خلاصه کنیم. چنین عددی می‌تواند معیار مرکزی باشد. معیار مرکزی نحوه تمرکز داده‌ها را بیان می‌کند. مهمترین معیارهای مرکزی میانگین، میانه و نما است که در این بخش به شرح هر یک از آنها خواهیم پرداخت.
- میانگین همان معدل داده‌هاست. میانگین نمونه را با \bar{x} و میانگین جامعه را با μ نشان می‌دهند.



مثال:

اطلاعات قد ۱۶ مرد بالای ۲۰ سال در جدول ۵.۴ خلاصه شده است. میانگین قد مردان چقدر است؟

جدول ۵.۴ توزیع قد مردان بالای ۲۰ سال

کرانه های گروه	x_i	f_i	$f_i x_i$
۱۵۰-۱۶۰	۱۵۵	۲	۳۱۰
۱۶۰-۱۷۰	۱۶۵	۷	۱۱۵۵
۱۷۰-۱۸۰	۱۷۵	۶	۱۰۵۰
۱۸۰-۱۹۰	۱۸۵	۱	۱۸۵
جمع		۱۶	۲۷۰۰

$$\mu = \frac{\sum_{i=1}^j f_i x_i}{N} = \frac{۲۷۰۰}{۱۶} = ۱۶۸/۷۵ \text{ سانتیمتر}$$

واحد رصفهان خوراکان

(Median میانه)

داده ای است که نیمی از داده ها از آن بزرگتر و نیمی دیگر از آن کوچکتر هستند. به عبارت دیگر تمامی داده ها را به دو قسمت مساوی تقسیم می کند.

اگر داده ها را از کوچک به بزرگ مرتب نماییم، عدد m را میانه این داده ها می نامیم، اگر نصف داده ها در سمت چپ و نصف داده در سمت راست این عدد قرار گیرد.

(Mode مد یا نما)

داده ای که فراوانی آن نسبت به دیگر داده ها بیشتر باشد، نما یا مد نامیده می شود و آن را با نماد M نمایش می دهیم.

معیارهای پراکندگی

با وجود این که در بسیاری از موارد، میانگین توصیف نسبتاً کاملی از مجموعه داده‌ها ارائه می‌دهد، اما گاهی وجود اطلاعات بیشتر در مورد داده‌ها ضروری است. یک مفهوم مهم در ارتباط با داده‌های آماری، **میزان تغییرات** آنهاست، بدین معنی که اندازه‌گیریها تا چه اندازه از فردی به فرد دیگر یا شیی به شیی دیگر تغییر می‌کنند. در این بخش، به بررسی و محاسبه میزان تغییرات به عنوان معیارهای پراکندگی خواهیم پرداخت. مهمترین معیارهای پراکندگی عبارتند از **دامنه، میانگین انحراف ها از میانگین یا از میانه، واریانس و انحراف معیار است.** علاوه بر مطالب فوق، در این بخش داده‌های استاندارد و ضریب تغییرات را نیز معرفی خواهیم کرد.

واریانس (پراش) و انحراف از معیار بهترین شاخصها جهت بررسی پراکندگی داده‌های کمی هستند.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

واریانس جامعه

در حالت طبقه بندی

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^k f_i (x_i - \mu)^2$$

$$\sigma^2 = \frac{\sum_{i=1}^n f_i (x_i - \mu)^2}{N} = \frac{\sum_{i=1}^n f_i x_i^2 - \frac{(\sum_{i=1}^n f_i x_i)^2}{N}}{N}$$

واریانس نمونه

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

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

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^k f_i x_i^2 - \frac{1}{n} \left(\sum_{i=1}^k f_i x_i \right)^2 \right]$$

در حالت طبقه بندی

انحراف از معیار

جذر واریانس را انحراف از معیار گوئیم.

ضریب تغییرات

ضریب تغییر عبارت است از اندازه نسبی انحراف معیار در مقایسه با میانگین. برخلاف واریانس و انحراف از معیار ضریب تغییر به واحد اندازه‌گیری وابسته نیست و برای مقایسه جمعیت‌های یکسان به کار می‌رود. در مقایسه هر اندازه که ضریب تغییر ویژگی جمعیتی کمتر باشد، ویژگی آن جمعیت بهتر ارزیابی می‌شود.

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

مثال



واریانس و انحراف معیار قد مردان را در مثال قبل محاسبه کنید.

کرانه های گروه	x_i	f_i	x_i^2	$f_i x_i^2$	$f_i x_i$
۱۵۰-۱۶۰	۱۵۵	۲	۲۴۰۲۵	۴۸۰۵۰	۳۱۰
۱۶۰-۱۷۰	۱۶۵	۷	۲۷۲۲۵	۱۹۰۵۷۵	۱۱۵۵
۱۷۰-۱۸۰	۱۷۵	۶	۳۰۶۲۵	۱۸۳۷۵۰	۱۰۵۰
۱۸۰-۱۹۰	۱۸۵	۱	۳۴۲۲۵	۳۴۲۲۵	۱۸۵
جمع		۱۶		۴۵۶۶۰۰	۲۷۰۰

$$\sigma^2 = \frac{\sum_{i=1}^n f_i x_i^2 - \frac{(\sum_{i=1}^n f_i x_i)^2}{N}}{N}$$

$$\sigma^2 = \frac{456600 - \frac{(2700)^2}{16}}{16} = 61$$

انحراف معیار قد برابر خواهد بود با:

$$SD = \sqrt{61} = 7.8 \text{ cm}$$

مدرس: دکتر سمیه عباسی

عنوان درس: تحلیل آماری-فصل اول: آمار توصیفی



داده های زیر توزیع دیابت خون در ۲۲ نفر از افراد بالای ۵۰ سال را نشان می دهد. مطلوبست محاسبه میانگین، میانه، مد، واریانس، انحراف از معیار و ضریب تغییرات

کرانه های گروه	f_i	x_i
۱۴۰-۱۵۰	۱	۱۴۵
۱۵۰-۱۶۰	۱۳	۱۵۵
۱۶۰-۱۷۰	۵	۱۶۵
۱۷۰-۱۸۰	۳	۱۷۵
جمع	۲۲	

داده های استاندارد

فرض کنیم میانگین و انحراف معیار داده ها بترتیب \bar{x} و s باشد. اگر از هر داده x را کم و بر s تقسیم کنیم. داده ها بفرم استاندارد تبدیل می شوند.

داده های استاندارد

$$z_i = \frac{x_i - \bar{x}}{s}, \quad i=1, \dots, k$$

به سادگی می توان نشان داد که داده های استاندارد دارای میانگین برابر با صفر و واریانس برابر با یک هستند و به واحد اندازه گیری بستگی ندارند.

مثال

- فرض کنید یک دسته از دانشجویان در دو امتحان شرکت کرده‌اند و خلاصه نتایج آزمونها به شرح زیر است.
- آزمون اول: میانگین نمرات برابر ۶۰، انحراف معیار برابر ۶، ماکزیمم نمره از ۱۰۰
- آزمون دوم: میانگین نمرات ۷۰، انحراف معیار برابر ۷، ماکزیمم نمره از ۱۰۰۰
- الف- چگونه این دو نتیجه را با هم مقایسه و ارزیابی می‌کنید؟
- ب- اگر دانشجویی در آزمون اول نمره ۶۵ و در آزمون دوم نمره ۷۲۰ را کسب کرده باشد، وضعیت دانشجو در کدام آزمون مطلوبتر است؟
- حل: الف) با محاسبه ضریب تغییر دو آزمون معلوم می‌شود که
- هر چه ضریب تغییرات کوچکتر باشد، بهتر است.

$$v_1 = \frac{s_1}{x_1} = \frac{6}{60} = \frac{1}{10} = 10\%$$

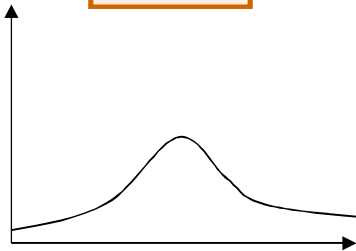
$$v_2 = \frac{s_2}{x_2} = \frac{7}{700} = \frac{1}{100} = 1\%$$

حل: ب) با محاسبه نمرات استاندارد:

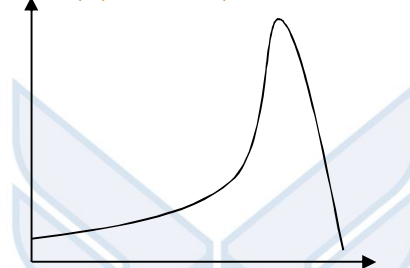
$$z_1 = \frac{65 - 60}{6} = \frac{5}{6}, \quad z_2 = \frac{720 - 700}{7} = \frac{20}{7}$$

منحنیهای فراوانی:

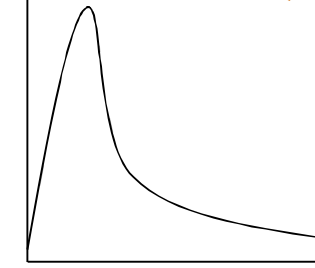
مقارن



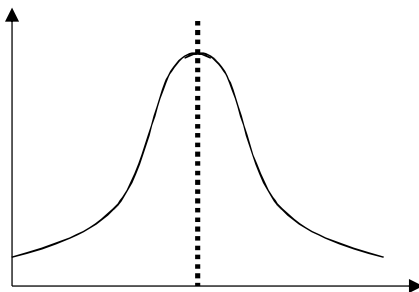
نامتقارن / چوله به چپ



نامتقارن / چوله به راست

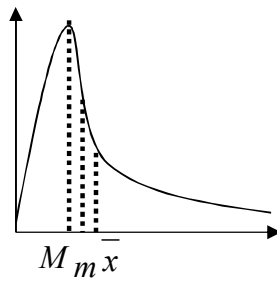


منحنی نرمال استاندارد



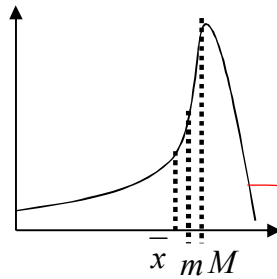
منحنیهای فراوانی در طبیعت تنوع زیادی دارند، اما بسیاری از منحنیهای فراوانی تک نمایی یا مقارن هستند یا چوله و یا برجسته و یا پخ. ایده آلترین منحنی فراوانی مقارن، منحنی فراوانی نرمال استاندارد است. برای منحنیهای فراوانی کاملاً مقارن تک نمایی مقادیر میانگین، میانه و نما بر هم منطبق میشوند. در طبیعت، عموماً منحنی فراوانی مقارن ایده آل کمتر یافت میشود و بسیاری از منحنیهای فراوانی موجود در طبیعت نامتقارن برجسته یا پخ هستند. میزان انحراف از تقارن ایده آل را معمولاً با دو معیار چولگی و برجستگی می‌سنجند.

چولگی



$$M < m < \bar{x}$$

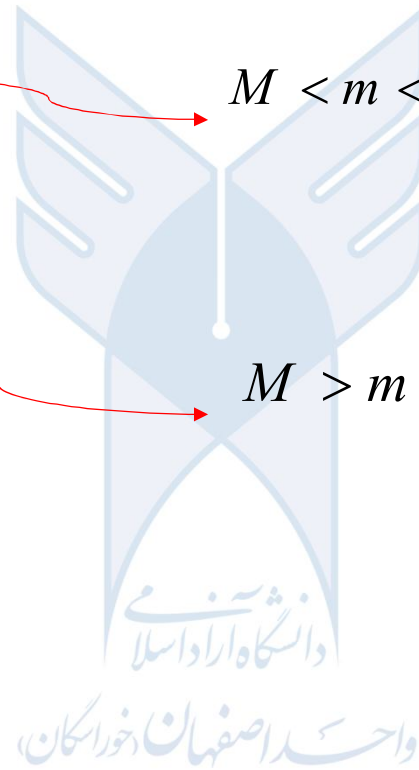
به راست یا مثبت



$$M > m > \bar{x}$$

به چپ یا منفی

چولگی:



انواع متغیرهای پژوهش

- متغیرهای مستقل و وابسته
- متغیرهای کمی و کیفی
- متغیرهای فعال و خصیصه ای
- متغیرهای دو ارزشی و چند ارزشی

متغیرهای مستقل و وابسته

متغیرها بر اساس نقشی که در پژوهش ایفا می کنند به دو دسته مستقل و وابسته تقسیم می شوند. متغیر مستقل به متغیری گفته می شود که از طریق آن متغیر وابسته، تبیین یا پیش بینی می شود. یعنی متغیر مستقل توسط پژوهشگر انتخاب می شود تا تأثیر یا رابطه‌ی آن با متغیر دیگر اندازه گیری شود. در تحقیقات آزمایشی، متغیر مستقل توسط پژوهشگر دست کاری می شود تا تأثیر آن بر متغیر وابسته مشخص شود اما در تحقیقات غیرآزمایشی متغیر مستقل توسط پژوهشگر دست کاری نمی شود و در این حالت متغیری است که از پیش فرض شده است.

متغیرهای مستقل و وابسته

متغیر وابسته، متغیری است که مشاهده یا اندازه گیری می شود تا تأثیر متغیر مستقل بر آن مشخص شود. وجود متغیر وابسته بستگی به متغیر مستقل دارد. در مطالعات از نوع همبستگی، به جای متغیر مستقل از اصطلاح متغیر پیش بینی و به جای متغیر وابسته از متغیر ملاک استفاده می شود.

برای تشخیص متغیر مستقل از متغیر وابسته باید این قاعده ی کلی را همیشه در نظر گرفت که متغیر مستقل، آزمودنی ها را گمارش نماید. یعنی یک ویژگی یا خصوصیت است که بعد از انتخاب توسط محقق دست کاری می شود و مقادیری را می پذیرد تا تأثیر یا ارتباط آن بر روی متغیر وابسته مشاهده شود. به عنوان مثال، در یک طرح آزمایشی که به منظور تعیین کارایی یک روش تدریس جدید بر میزان یادگیری علوم دانش آموزان طرح ریزی شده است، روش تدریس جدید، متغیر مستقل است و یادگیری علوم توسط دانش آموزان، متغیر وابسته می باشد.

متغیرهای کمی و کیفی

به طور کلی متغیرها از نظر ماهیت و مقادیری که می‌پذیرند به دو دسته‌ی متغیرهای کمی و متغیرهای کیفی تقسیم می‌شوند.

متغیرهای کیفی را متغیرهای مقوله‌ای نیز می‌نامند که شامل حالت‌های گوناگون یک ویژگی است، مانند: جنسیت، شغل، مذهب، و غیره... پژوهشگر توان اندازه‌گیری متغیر کیفی را ندارد و ویژگی‌های آن را نمی‌تواند توسط ارقام ریاضی نمایش دهد. البته می‌توان اعداد را به عنوان اسم یا کد مورد استفاده قرار داد ولی ارزشی عدد ندارند. روشن است که متغیرهای کیفی را نمی‌توان جمع یا تفریق کرد و برای آنها مبدا اندازه‌گیری وجود ندارد. متغیرهای کیفی را می‌توان در برخی از موارد با مقیاس رتبه‌ای، اندازه‌گیری کرد.

متغیرهای کمی و کیفی

متغیرهای کیفی، متغیرهایی هستند که برای اندازه گیری آنها می توان اعداد را به دو وضعیت آزمودنی و بر طبق قاعده‌ای معین منتسب کرد. این گونه متغیرها به دو دسته متغیر کمی پیوسته و متغیر کمی گسسته تقسیم می شوند. متغیر کمی پیوسته، متغیری است که بین هر دو مقدار متوالی آن، مقادیر بی شماری وجود دارد. مانند بهره هوشی، قد و وزن. در حالی که بین هر دو مقدار متغیر کمی گسسته نمی توان مقدار دیگری را پیدا نمود. مانند تعداد فرزندان. به عبارت دیگر، متغیر کمی به متغیری اطلاق می شود که از نظر کمیت تغییر می کند و اختلاف مقادیر آن را می توان با استفاده از عدد ثبت کرد و آنها را می توان با هم جمع کرد.



متغیرهای فعال و خصیصه ای

به متغیرهایی که دست کاری می‌شوند و محقق می‌تواند آزمودنی‌ها را در سطح مختلف آن به صورت تصادفی جایگزین کند متغیرهای فعال می‌گویند. مثلاً وقتی پژوهشگری روش‌های مختلف تدریس را به کار می‌برد، یا به یک گروه از کودکان پاداش داده و گروه دیگری را تنبیه می‌کند، متغیرهای روش تدریس و تقویت (تشویق، تنبیه) را به گونه‌ای فعال دست کاری می‌کند. این متغیرها، مشابه متغیر مستقل می‌باشند.

متغیرهای خصیصه ای به متغیرهایی که اطلاق می‌شود که بدون دخالت محقق، قبلاً در سطوح مختلف جایگزین شده‌اند. همه‌ی متغیرهایی که نشانگر ویژگی‌های آدمی است و پژوهشگر نمی‌تواند آنها را تغییر دهد متغیر خصیصه ای هستند. بنابراین متغیرهایی نظیر هوش، استعداد، جنسیت، موقعیت اجتماعی-اقتصادی، متغیرهای خصیصه ای هستند و توسط پژوهشگر قابل دست کاری نیستند. به همین دلیل این متغیرها را گاهی متغیرهای ارگانیزمی نیز می‌نامند.

متغیرهای خصیصه ای در بعضی از موارد همان متغیرهای کیفی هستند.

دانشگاه آزاد اسلامی

نکته

متغیر خصیصه ای گاهی اوقات می‌تواند در نقش متغیر فعال قرار گیرد. مثلاً جنسیت با اینکه یک متغیر خصیصه ای است ولی بعضی اوقات می‌تواند از طرف محقق کنترل شود تا تأثیر آن بر متغیر دیگر بررسی شود.

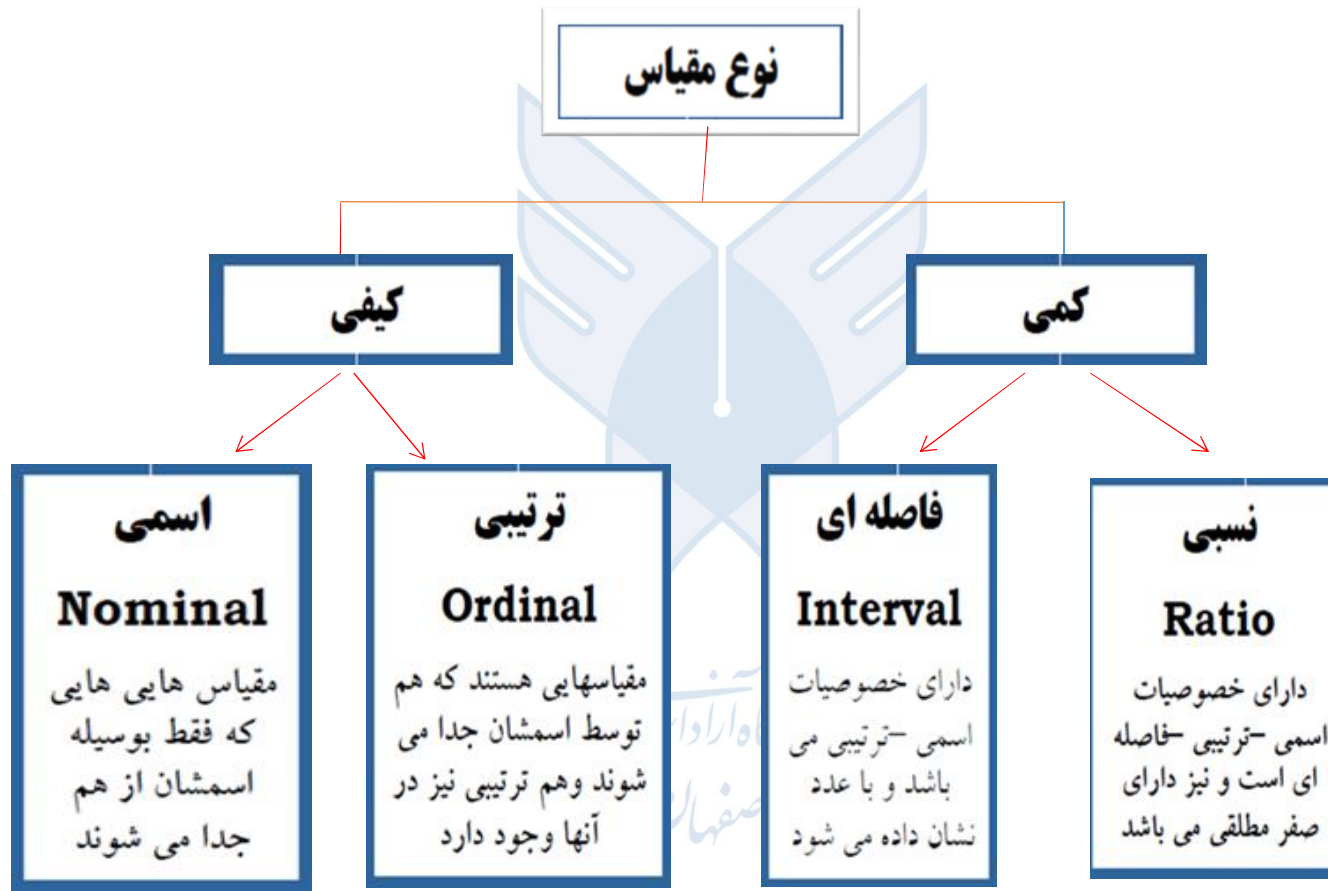
متغیرهای دوارزشی و چند ارزشی

متغیر دوارزشی به متغیری اطلاق می‌شود که به آن فقط دو ارزش یا دو عدد نسبت داده می‌شود، مانند جنس که دارای دو ارزش زن و مرد است و می‌توان برای ثبت آنها از اعداد صفر و یک یا اعداد دیگری استفاده کرد.

کرلینجر (۱۹۸۶) متغیر دوارزشی را دوبخشی نامیده است و چنین می‌نویسد: «برخی از این متغیرها، دو بخشی واقعی هستند مانند زن و مرد، مرگ و حیات و شهری و روستایی که حضور یا عدم حضور ویژگی موجب تقسیم بندی آنها می‌شود و برخی از آنها ممکن است دوبخشی ساختگی باشند».

متغیر چندارزشی متغیری است که بیش از دو عدد یا دوارزش به آن اختصاص داده می‌شود، مانند سطح تحصیل و هوش که دارای درجات مختلفی هستند و به هر یک از درجات آنها می‌توان عدد یا ارزش معینی را اختصاص داد.

مقیاس‌های اصلی اندازه‌گیری



مقیاس
اسمی

مربوط به **صفات** است و شامل طبقه های **نا منظم و مانعه الجمع** می باشد.
در مقیاس اسمی هیچ نوع ترتیب یا شدت وضعی وجود ندارد. **پایین ترین سطح اندازه گیری است.** و وجود یا عدم وجود یک صفت را نشان می دهد. مانند جنسیت ، مذهب ، رشته تحصیلی

عملیات مجاز آماری : **فراوانی ، نما**

عملیات مجاز ریاضی : هیچ عملیاتی مجاز نیست.

مقیاس
ترتیبی

موقعیت فرد یا شیء بر حسب **صفت معین** در گروه با **اعداد** مشخص میشود و گروه از **بالاترین به پایین تر** مرتب می شود.

اختلاف رتبه ها کمی نیست بنابراین **میانگین معنا ندارد** بهترین آماره توصیفی ، **میانها** است مثل طبقه اقتصادی و اجتماعی.

عملیات مجاز آماری : **فراوانی ، نما ، میانها ، محاسبه درصد و محاسبه ضریب همبستگی رتبه ای اسپیرمن.**

عملیات مجاز ریاضی : هیچ عملیاتی مجاز نیست.

مقیاس
فاصله ای

طبقه بندی، نام گذاری، مرتب کردن و مشخص کردن فاصله دقیق بین داده ها امکان پذیر است. فاقد صفر مطلق هستند و بر حسب صفر قراردادی روی یک پیوستار قرار میگیرند. مثل اندازه گیری دما

چون فواصل بین واحدها برابر است یکی از ویژگیهای مقیاس خوب را دارد. عملیات مجاز آماری: همه (فراوانی، نما، میان، میانگین، محاسبه درصد، واریانس، انحراف معیار، ضریب همبستگی پیرسون و ضریب همبستگی رتبه ای) عملیات مجاز ریاضی: جمع و تفریق

مقیاس
نسبی

علاوه بر همه ویژگی های مقیاس فاصله ای دارای صفر مطلق نیز هست مثل قد، وزن مقیاس نسبی قویترین مقیاس اندازه گیری است عملیات مجاز آماری و ریاضی: همه عملیات ها

نوع	آزمون‌های آماری	شاخص آماری	مثال	مقیاس‌ها
کیفی	ناپارامتریک	نما	مرد- زن	اسمی
کیفی	ناپارامتریک	مد و میانه	خیلی خوب، خوب؛ متوسط و ...	ترتیبی
کمی	پارامتریک ناپارامتریک	مد و میانه و میانگین	درجه سانتی گراد- ضرب هوشی	فاصله ای
کمی	پارامتریک ناپارامتریک	مد و میانه و میانگین	متر	نسبتی



GOOD LUCK