دانشگاه علوم پزشکی و خدمات بهداشتی درمانی بیرجند

**شبکه بهداشت و درمان شهرستان قاینات**

**مرکز بهداشت شهرستان قاینات**

**مرکز آموزش بهورزی و ماما روستایی قاینات**

# آشنایی با روش های آماری

*تهیه و تدوین: حجت کریمی- کارشناس آمار مرکز بهداشت شهرستان قاینات*

# بخش اول – مفاهیم اولیه

## ۱ – علم آمار –مفاهیم آماری–جمع آوری داده ها

**تعریف علم آمار:** علم آمار به مجموعه روشهای علمی اطلاق می شود که برای جمع آوری اطلاعات اولیه، مرتب و خلاصه کردن، طبقه بندی و تجزیه و تحلیل اطلاعات اولیه و تفسیر آنها بکار می رود .

**تعریف جامعه آماری:** هر مجموعه ای از ااشیاء یا افرادی که لاقل دارای یک صفت مشترک باشد را جامعه آماری می گویند .

- چرا با جامعه آماری سرو کار پیدا می کنیم ؟ چون اگر تک تک اشیاء یا افراد را مطالعه کنیم ممکن است اطلاعی بدست نیاوریم ولی وقتی افراد را به صورت مجموعه در نظر می گیریم خواصی در آن ظاهر می شود مثلاً در یک کشور رشد جمعیت سالانه ۳٪ می باشد ولی مطالعه تک تک افراد جامعه این اطلاعات را به ما نمی دهد .
- هر یک از اشیاء یک جامعه آماری را یک فرد جامعه می نامیم. وقتی می گوییم دستمزد های یک کارخانه، دستمزد نیز یک فرد جامعه است.( فرد لازم نیست انسان باشد.) مجموع افراد یک جامعه را حجم جامعه می نامیم و آن را با $N$ نشان می دهیم.

**تعریف متغیر:** صفاتی از هریک از افراد جامعه آماری که از یک فرد به فرد دیگر تغییر می کنند متغیر می باشند. مثلاً در جامعه دانشجویان دانشگاه تهران،صفاتی مانند جنس، وزن، قد و غیره متغیرهای جامعه می باشند.متغیرها به دو دسته تقسیم می شوند:

۱)متغیر کیفی: متغیرهایی هستند که واحد ندارند و قابل شمارش و یا اندازه گیری نیستند ولی می توان آنها را طبقه بندی کرد. مانند جنس، شغل و نوع بیماری. متغیر های کیفی نیز دو قسمند:

الف) متغیر اسمی: متغیرهای کیفی که قابل مقایسه با همدیگر نیستند مانند رنگ چشم که مثلاً مشکی یا میشی نمی توان گفت که مشکی از میشی بهتر است، یا در مورد اهل شهر تهران با مشهد بودن نمی توان اهل تهران بودن را با اهل مشهد بودن مقایسه کرد. این متغیرها فقط اسم هستند .

ب) متغیر ترتیبی: متغیرهای کیفی که شدت و ضعف را نشان می دهند یعنی ترتیب بین اعداد رعایت شده است. مثلا می خواهیم کیفیت ۵ نوع پارچه را بررسی کنیم. ما آنها را ردیف می کنیم و کد ۱و۲و۳و۴و۵ را به آنها می دهیم در اینجا کیفیت ۲ دو برابر کیفیت ۱ نیست ولی می توان گفت که کیفیت ۲ از ۱ بهتر است .

۲)متغیر های کمی: متغیرهایی هستند که قابل اندازه گیری یا شمارش و یا قابل مقایسه و سنجش هستند. این متغیرها نیر بر دو قسمند:

الف) متغیرهای کمی گسسته: متغیرهای قابل شمارش هستند که بین مقادیر قابل تصور از آن فاصله وجود داشته باشد. مانند تعداد افراد خانوار.

ب) متغیرهای کمی پیوسته : متغیرهای کمی هستند که مقادیر خودشان را از اعداد حقیقی می گیرند یعنی فاصله ای بین هیچ یک از دو مقدار قابل تصور از متغیر وجود ندارد . مانند قد و وزن یا طول.

**بررسی آماری:** بررسیی است که موضوع مورد مطالعه را به یک جامعه مربوط می کند و در آن جامعه افراد را مورد مطالعه قرار می دهد. بررسی های آماری شامل سه مرحله است .

۱) مشاهده

۲) گروه بندی تهیه جداول و رسم نمودارها

۳) محاسبه شاخص ها (مشخص کننده ها) و تجزیه و تحلیل آنها

**آمارگیری:** در مطالعات آماری که اطلاعات آماری را نتوان از ثبت جاری بدست آورد، از طریق آمارگیری استفاده می کنیم. مشاهدات بطور کلی خود بر دو نوع هستند :

۱) سراسری: کلیه افر اد جامعه را مورد مطالعه قرار می دهیم و معمولاًاین نوع مشاهدات را سرشماری می نامیم و اصولاً سرشماری خاص انسان است. ولی امروز در تمام زمینه ها بکار می رود .

۲) غیر سراسری: مشاهداتی هستند که در تمام افراد جامعه مورد مطالعه قرار نمی گیرند و خود به چند دسته تقسیم می شوند:

الف) آمار گیری نمونه ای : در زیر چند روش نمونه گیری به طور فهرست وار اشاره می گردد.

- نمونه گیری تصادفی : یکی از دقیق ترین روشهای آمار گیری است که افراد مورد مطالعه بطور تصادفی و بر طبق قانون احتمالات انتخاب می شوند. بطوری که این جامعه نمونه نماینده جامعه اصلی باشد. هر نمونه که با یک شانس معلوم انتخاب شده باشد نمونه تصادفی نامیده می شود . اگر روند انتخاب نمونه طوی باشد که شانس انتخاب برای هر نمونه ممکن (با حجم ثابت از همان جامعه ) برابر باشد آن را نمونه تضادفی ساده می نامیم ولی اگر شانس انتخاب هر یک از اعضای نمونه برابر نباشد آن را نمونه تصادفی ای با احتمال متغیر می نامیم نمونه برداری تصادفی ساده ای را می توان به روش با جایگذاری و یا بدون جایگذاری انجام داد.

- نمونه گیری خوشه ای : یک نمونه گیری تصادفی ساده است که به جای یک فرد گروههایی از افراد جامعه به عنوان واحد انتخابی در نظر گرفته می شوند و آن را می توان با جایگذاری و یا بدون جایگذاری اجرا کرد .

- نمونه گیری تصادفی طبقه ای : فرض کنید جامعه را به $k$ طبقه متساوی الحجم تقسیم کنیم و بخواهیم از جامعه یک نمونه $n$ تایی انتخاب کنیم. می توان از هر یک از طبقات یک نمونه به اندازه $\frac{n}{k}$ انتخاب کرد . حال اگر طبقه ها هم حجم نباشد، نسبت نمونه ی متشابه نمی شود. یعنی متغیر می باشد. در این صورت آن را نمونه گیری تصادفی با احتمال متغیر می نامیم.

ب) آمار گیری با روش توده اصلی : در این روش تمام افراد جامعه را مورد مطالعه قرار نمی دهیم. بلکه جزئی از جامعه که سهم همه موضوع مورد مطالعه را در بردارند را مورد توجه قرار می دهیم.

ج) آمار گیری با روش یکه نگاری : در این روش به جای آنکه تمام واحد های یک جامعه را مطالعه کنیم و یا تعدادی را بر اساس روشهای تصادفی انتخاب کنیم فقط یک واحد جامعه را مطالعه می کنیم و در آن واحد به جزئیات می پردازیم که البته در جای خود با ارزش است ولی از لحاظ تعمیم به کل جامعه بی ارزش می باشد.

د ) آمار گیری با روش مکاتبه : در این روش ما یک پرسش نامه تنظیم می کنیم و آن را برای افراد جامعه می فرستیم و جوابهای رسیده را مطالعه می کنیم.

## ۲- توزیعهای فراوانی

الف) توزیع فراوانی داده های کیفی : فرض کنید $X$ یک صفت کیفی با حالات $E_i \ ,i = 1,2,3,\dots,k$ باشد. مثلاً $X$ صفت کیفی گروه خونی افراد با حالات $A,B,O,AB$ باشد . همچنین فرض کنید $f_1$ تعداد افراد دارای حالت $E_1$ و $f_2$ تعداد افراد دارای حالت $E_2$ و ... و $f_k$ تعداد افراد دارای حالت $E_k$ باشد. اگر حجم نمونه انتخابی برابر $n$ باشد داریم:

$$\sum_{i=1}^{k} f_i = f_1 + f_2 + \dots + f_k = n$$

حال اگر جدولی مطابق زیر تشکیل از دو ستون که یکی نشان دهنده حالات و دیگری نشان دهنده تعدادافرادی است که دارای آن حالت هستند را تشکیل دهیم ، چنین جدولی را جدول توزیع فراوانی می نامیم .

| حالات | $f_i$ |
|---|---|
| $E_1$ | $f_1$ |
| $E_2$ | $f_2$ |
| . | . |
| $E_k$ | $f_k$ |

مثال۱- کمیته پزشکی متشکل از ۵ پزشک ۳ پرستار و ۲ بهیار می باشد. جدول توزیع فراوانی آن را رسم کنید.

حل :

| fi | حالات |
|---|---|
| ۵ | پزشک |
| ۳ | پرستار |
| ۲ | بهیار |
| ۱۰ | |

ب ) توزیع فراوانی داده های کمی: که دو نوع می باشد، داده های کمی گسسته و پیوسته. با ۲ مثال این دو نوع را توضیح می دهیم .

مثال۲- فرض کنید می خواهیم۲۰ دانشجو را از نظر تعداد افراد خانوار که یک صفت کمی گسسته است طبقه بندی کنیم. فرض کنید تعداد خانوار ۲ نفره، ۳ و تعداد خانوار ۳ نفره، ۵ و تعداد خانوار ۴ نفره ۶۰۰ و تعداد خانوار ۵ نفره ۵۰ و تعداد خانوار ۶ نفره،۱ نفر می باشد.جدول توزیع فراوانی آن را رسم کنید.

حل:

| $f_i$ | $x_i$ |
|---|---|
| ۳ | ۲ |
| ۵ | ۳ |
| ۶ | ۴ |
| ۵ | ۵ |
| ۱ | ۶ |
| ۲۰ | |

حال اگر صفت مورد بررسی صفت پیوسته باشد مانند وزن افراد، طول قد و غیره، نمی توان برای هرنقطه فراوانی تعیین نمود. مثلا بگوییم ۴ نفر دارای وزن ۶۵ کیلو گرم هستند زیرا وزن این افراد درفاصله (۶۵/۵-۶۴/۵ ) است.لذا درمورد داده های پیوسته مجاز هستیم برای طبقات فاصله ای را منظورکنیم برای طبقه بندی مراحل زیررا انجام می دهیم:

۱) مرتب کردن داده ها از کوچک به بزرگ
۲) تعین دامنه
۳) تعداد طبقات
۴) فاصله طبقات
۵) تکمیل جدول فراوانی

برای پیاده سازی این مراحل مثال می زنیم.

مثال ۳ - فرض کنید داده های زیر قد ۲۵ کودک مهد را بر حسب سانتیمتر نشان می دهد: -۱۰۵-۱۰۴-۱۱۰-۱۰۳
۱۰۵-۱۰۹-۹۸-۹۹-۱۰۴-۱۰۳-۱۰۹ -۱۰۰-۱۰۵-۹۸-۱۰۳-۱۰۴-۱۱۰-۱۰۶-۱۰۳-۱۱۰-۱۰۲-۱۰۴-۹۷-۱۰۱-۱۰۰

جدول توزیع فراوانی آن را رسم کنید.

**توجه:** اعداد فوق گرد شده اند.مثلا ۱۱۰/۳ را ، ۱۱۰و۹۸/۷ را با ۹۹ نشان می دهیم.

حل : ابتدا داده ها را از کوچک به بزرگ مرتب می کنیم:

۱۰۹-۱۰۶-۱۰۵-۱۰۵-۱۰۵-۱۰۴- ۱۰۴-۱۰۴-۱۰۴-۱۰۴-۱۰۳-۱۰۳-۱۰۳-۱۰۳-۱۰۲-۱۰۱-۱۰۰-۱۰۰-۹۹-۹۸-۹۷
-۱۰۹-۱۱۰-۱۱۰-۱۱۰

اکنون با یک نگاه اجمالی می توان اطلاعاتی را کسب نمود. مثلاً قد کوتاهترین کودک ۹۷ و بلندترین آن ۱۱۰ ساتیمتر است و یا قد بیش از۵۰٪ کودکان بین۱۰۰تا۱۰۶سانتیمتر است.هرچند این اطلاعات کافی نیست ولی برای شروع مفید است.

حال دامنه یا دامنه تغییرات را پیدا می کنیم . دامنه عبارت است از تفاضل بزرگترین داده از کوچکترین داده و آن را با حرف R نشان می دهیم: R = بزرگترین داده – کوچکترین داده = ۱۱۰– ۹۷+۱ =۱۴

توجه کنید که کمترین داده می تواند کمی بیشتر از ۹۶/۵ و بیشترین داده می تواند کمتر از ۱۱۰/۵ باشد در واقع دامنه تغییرات (۱۱۰/۵–۹۶/۵) است.

حال برای طبقه بندی و تعیین طبقات معمولاً تعداد طبقات و بین ۵ تا ۲۰ انتخاب می کنند . البته تعداد طبقات باید طوری انتخاب شود که اطلاعات قابل مطلاحظه ای از دست نرود.

فاصله طبقات از تقسیم بر تعداد طبقات با تقریبی اضافی به دست می آید.

$$C = \text{فاصله طبقات} = \frac{\text{دامنه}}{\text{تعداد طبقات}} = \frac{R}{\text{تعداد طبقات}}$$

مثلاً اگر تعداد طبقات را ۵ اختیار کنیم در این صورت فاصله طبقات برابر است با:

$$C = \frac{14}{5} = 2.8 \cong 3$$

حال می خواهیم جدول توزیع فراوانی را برای این مثال تشکیل دهیم .

**فروانی** : فروانی هر طبقه یعنی تعداد دفعات تکرار داده ها در آن و آن را با نماد $f$ نشان می دهیم.

– در ستون چهارم فراوانی هر طبقه را با عدد می نویسیم .

– خارج قسمت فراوانی هر طبقه بر کل فراوانی را فراوانی نسبی می نامیم و آن را با $f_p$ نشان می دهیم

– ۱۰۰ برابر فراوانی نسبی هر طبقه را در حد فراوانی نسبی آن طبقه می نامیم و آن را با نماد $P$ نشان می دهیم .

– فراوانی تجمعی هر طبقه عبارت است از مجموع فراوانی های آن طبقه بالاتر و آن را با نماد $F_c$ نشان می دهیم .

– فراوانی تجمعی نسبی هر طبقه عبارت است از مجموع فراوانی های نسبی آن طبقه با طبقات بالاتر و آن را با نماد $F_{cp}$ نشان می دهیم .

– در صد فراوانی تجمعی نسبی هر طبقه یعنی ۱۰۰ برابر فراوانی تجمعی نسبی آن طبقه و آن را با نماد $P_c$ نشان می دهیم .

– نماینده هر طبقه عبارت است از میانگین کران بالا و کران پائین آن طبقه

$x_i$ =۰/۵(کران بالا+کران پایین )

مطالب بیان شده را در مورد مثال مورد بحث در جدول زیر ملاحظه می کنید .

| $P_c$ | $F_{cpi}$ | $F_{ci}$ | $P_i$ | $f_{pi}$ | $f_i$ | $x_i$ | حدود طبقات | حدود واقعی |
|---|---|---|---|---|---|---|---|---|
| ۱۶ | ./۱۶ | ۴ | ۱۶ | ۰/۱۶ | ۴ | ۹۸ | ۹۷–۹۹ | ۹۶/۵–۹۹/۵ |
| ۳۲ | ۰/۳۲ | ۸ | ۱۶ | ۰/۱۶ | ۴ | ۱۰۱ | ۱۰۰–۱۰۲ | ۹۹/۵–۱۰۲/۵ |
| ۷۶ | ۰/۷۶ | ۱۹ | ۴۴ | ۰/۴۴ | ۱۱ | ۱۰۴ | ۱۰۳–۱۰۵ | ۱۰۲/۵–۱۰۵/۵ |
| ۸۰ | ۰/۸۰ | ۲۰ | ۴ | ٪۴ | ۱ | ۱۰۷ | ۱۰۶–۱۰۸ | ۱۰۵/۵–۱۰۸/۵ |
| ۱۰۰ | ۱ | ۲۵ | ۲۰ | ٪۲۰ | ۵ | ۱۱۰ | ۱۰۹–۱۱۱ | ۱۰۸/۵–۱۱۱/۵ |
| - | – | – | ۱۰۰ | ۱ | ۲۵ | | | |

جدول (۱)

از کاربردهای مهم این جدول این است که با یک نگاه به این جدول مثلاً ملاحظه می کنیم قد۴۴٪ کودکان در فاصله ( ۱۰۵/۵ و ۱۰۲/۵ ) سانتیمتر می باشد و یا قد ۷۶٪ کودکان کمتر از ۱۰۵/۵ سانتیمتر است.

– اگر داده های گسسته با تعداد کم داشته باشیم مانند مثال زیر لزومی ندارد که آنها را طبقه بندی کنیم ولی می توان جدولی مشابه با جدول (۱) برای آن تشکیل داد.

مثال ۴–: جدول توزیع فراوانی مربوط به تعداد فرزندان دختر در ۱۰۰۰ خانوار بصورت زیر بیان شده است جدولی مشابه جدول (۱) برای آن تشکیل دهید .

| تعداد خانواده | تعداد فرزندان دختر |
|---|---|
| ۴۰ | ۰ |
| ۱۴۰ | ۱ |
| ۳۴۰ | ۲ |
| ۲۹۰ | ۳ |
| ۱۶۵ | ۴ |
| ۲۵ | ۵ |
| ۱۰۰۰ | |

حل:

| $x_i$ | $f_i$ | $f_{pi}$ | $P_i$ | $F_{ci}$ | $F_{cpi}$ | $P_c$ |
|---|---|---|---|---|---|---|
| ۰ | ۴۰ | ٪۴ | ۴ | ۴۰ | ٪۴ | ۴ |
| ۱ | ۱۴۰ | ٪۱۴ | ۱۴ | ۱۸۰ | ٪۱۸ | ۱۸ |
| ۲ | ۳۴۰ | ٪۳۴ | ۳۴ | ۵۲۰ | ٪۵۲ | ۵۲ |
| ۳ | ۲۹۰ | ٪۲۹ | ۲۹ | ۸۱۰ | ٪۸۱ | ۸۰ |
| ۴ | ۱۶۵ | ٪۱۶۵ | ۱۶/۵ | ۹۷۵ | ٪۹۷۵ | ۹۷/۵ |
| ۵ | ۲۵ | ٪۲۵ | ۲/۵ | ۱۰۰۰ | ۱ | ۱۰۰ |

## ۳– بیان هندسی توزیع فراوانی ( نمودارهای آماری )

برای بهتر نمایاندن داده ها،معمولاً نمودار فراوانی داده ها را رسم می کنیم. در این بخش ابتدا نمودارهای فراوانی داده های کیفی و سپس نمودارهای فراوانی کمی را مورد بررسی قرار می دهیم .

الف) نمودار فراوانی داده های کیفی: می توان نمودارهایی برای داده های کیفی رسم نمود که در زیر در حل یک مثال به بیان آنها خواهیم پرداخت .

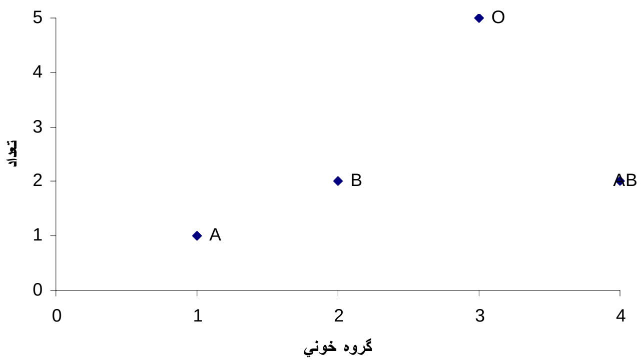
مثال ۵ – فرض کنید گروه خونی ۱۰ نفر تعیین و نتایج زیر بدست آمده است .

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

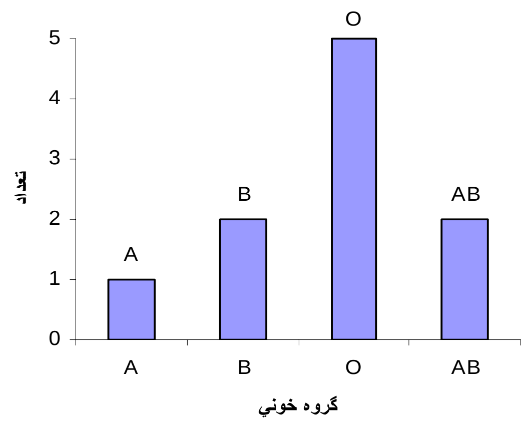
ابتدا نمایش جدولی ( جدول توزیع فراوانی ) را تشکیل می دهیم .

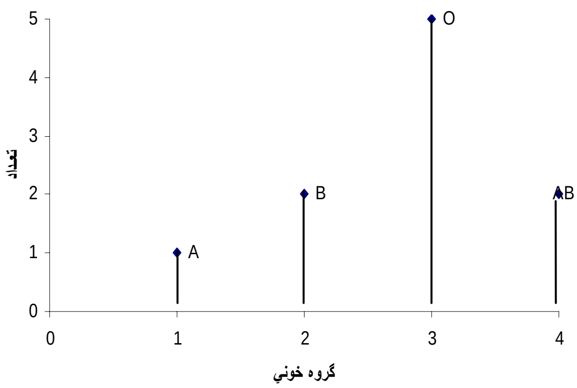
| $f_i$ | گروه خونی |
|---|---|
| ۱ | A |
| ۲ | B |
| ۵ | O |
| ۲ | AB |

حال دو محور عمود بر هم رسم می کنیم روی محور افقی چهار نقطه دلخواه برای AB,O,B,A انتخاب می کنیم. ( این محور مقیاس ندارد و ترتیب انتخاب نقاط مهم نیست زیرا داده ها از نوع کیفی- اسمی می باشد) روی محور عمودی تا بیشترین فراوانی مربوط به جدول توزیع فراوانی شماره گذاری می کنیم ( در مثال فوق این عدد ۵ می باشد) حال جدول توزیع فراوانی فوق را در محورهای عمده بر هم زیر پیاده سازی می کنیم .




حال نمودار فوق را نمودار سوزنی و اگر از این نقاط خطوطی بر محور افقی عمود کنیم آن را نمودار میله ای و اگر به جای خطوط در نمودار میله ای از نوارهای پهن استفاده ه شود آن را نمودار ستونی می نامیم. در زیر نمودارهای میله ای و ستونی را ملاحظه می کنید .



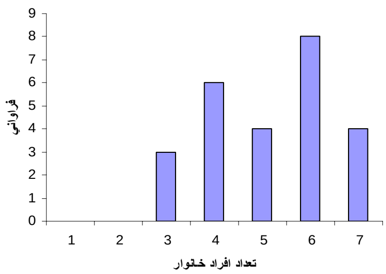



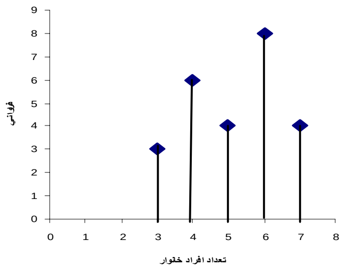

ب) نمودار فراوانی داده های کمی گسسته : اگر صفت متغیر کمی گسسته باشد، ابتدا دو محور عمود بر هم رسم می کنیم ورودی محور عمودی فراوانی ها ورودی محور افقی مقادیر متغیر را مشخص می کنیم هر مقدار و فراوانی آن یک زوج مرتب را تشکیل می دهند که یک نقطه را مشخص می کند مجموع این نقاط را نمودار توزیع فراوانی صفت کمی گسسته می نامیم .

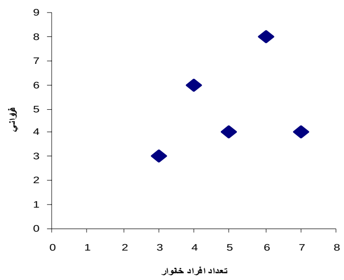
مثال ۶ – توزیع فراوانی افراد خانوار بصورت زیر مفروض است .

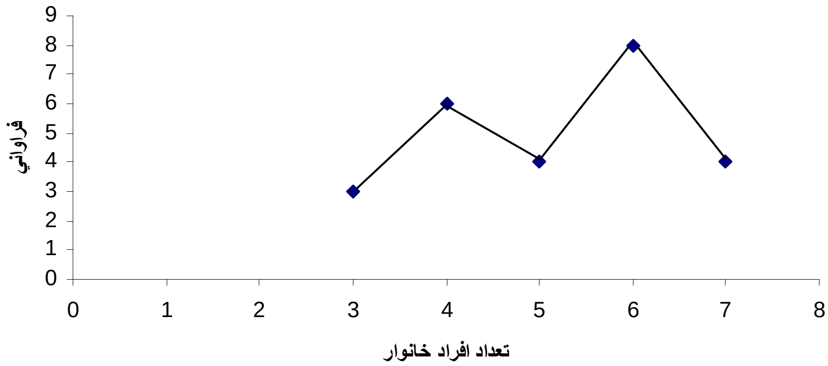
| $x_i$ | $f_i$ |
| --- | --- |
| ۳ | ۳ |
| ۴ | ۶ |
| ۵ | ۴ |
| ۶ | ۸ |
| ۷ | ۴ |
| | ۲۵ |

نمودارهای سوزنی، میله ای، ستونی و چند ضلعی فراوانی مثال فوق به شکل زیر است .








ج)نمودار فراوانی داده های کمی پیوسته : اگر صفت متغیر کمی پیوسته باشد برای بیان هندسی آن از هیستوگرام استفاده می شود. دو محور عمود بر هم رسم می کنیم و روی محور افقی حدود واقعی طبقات و روی محور عمودی فراوانی ها را مشخص می کنیم . فرض کنید می خواهیم هیستوگرام فراوانی جدول توزیع فراوانی (۱) (مثال طول قد کودکان ) را رسم کنیم. ابتدا عدد ۹۶/۵ را بطور دلخواه روی محور افقی اختیار کرده سپس عدد ۹۹/۵ را به دلخواه روی این محور انتخاب می کنیم حال محور افقی دارای مقیاس بوده و باید فاصله عدد ۱۰۳/۵ از ۹۹/۵ برابر فاصله قبلی باشد و به همین ترتیب ادامه می دهیم. آنگاه بیشترین فراوانی که در این مثال عدد ۱۱ است را روی محور عمودی به دلخواه انتخاب و سپس فاصله این عدد از مبدا را به ۱۱ قسمت مساوی تقسیم می کنیم و بدین ترتیب محور عمودی را مدرج می کنیم . حال مستطیلهایی به قاعده حدود واقعی فراوانی رسم می کنیم.

# بخش دوم- شاخصهای مرکزی و پراکندگی

## ۱- شاخصهای مرکزی

بطور کلی داده های یک جامعه آماری یک نوع تجمع و فشردگی در اطراف یک مقدار خاص از یک صفت متغیر را مورد مطالعه را بوجود می آورند . می خواهیم این مقدار خاص را به عنوان یک شاخص مرکزی مشخص کنیم و متذکر می شویم که یک شاخص مرکزی وقتی با ارزش است که دارای خواص زیر باشد :

۱ – در محاسبه آن از تمام داده ها استفاده شود.

۲ – دارای خصوصیات ساده قابل محاسبه باشد.

۳ – به فرم ریاضی قابل محاسبه باشد.

از شاخصهای مرکزی مهم به بررسی میانگین ، میانه و مد (نما) می پردازیم.

**الف ) میانگین**

میانگین خود بر چند نوع است که مهمترین آنها میانگین حسابی ، میانگین مهندسی ، میانگین وزنی ، میانگین همساز و میانگین درجه دوم می باشد. در اینجا ما فقط به میانگین حسابی می پردازیم.

میانگین حسابی : مهمترین شاخص مرکزی است و آن را به اختصار میانگین می نامیم. فرض کنید صفت متغیر دارای مقادیر $x_1, x_2, \dots, x_n$ باشد. اگر میانگین آنها را با $\bar{x}$ نشان دهیم:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

مثال ۱- داده های زیر نمرات دروس یک دانش آموز را نشان می دهد:

۲و۸و۴و۷و۶و۵و۹و۱و۳و۴

می خواهیم میانگین نمرات را حساب کنیم. داریم :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}(2+8+4+7+6+5+9+13+4) = \frac{49}{10} = 4.9$$

توجه۱- اگر توزیع فراوانی صفت گسسته x را به صورت زیر داشته باشیم:

| x | $x_1$ | $x_2$ | ... | $x_n$ |
|---|---|---|---|---|
| **فراوانی** | $f_1$ | $f_2$ | ... | $f_n$ |

با توجه به تعریف میانگین داریم:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i \quad , \quad \sum_{i=1}^{k} f_i = n \qquad (2)$$

مثال ۲ – جدول توزیع فراوانی نمرات ۲۵ دانش آموز بهورز به صورت زیر بیان شده است:

| $x_i$ | $f_i$ |
|---|---|
| ۱۴ | ۶ |
| ۱۵ | ۷ |
| ۱۷ | ۸ |
| ۱۹ | ۴ |
| | ۲۵ |

محاسبه میانگین به صورت زیر است:

$$\bar{x} = \frac{1}{25}\sum_{i=1}^{4} f_i x_i = \frac{1}{25}(6\times 14 + 7\times 15 + 8\times 17 + 4\times 19) = \frac{401}{25} = 16.04$$

توجه۲– اگر جدول توزیع فراوانی به صورت طبقه بندی شده باشد نماینده هر طبقه را در نظر می گیریم:

(کران پائین گروه iام + کران بالای گروه iام) $x_i = \frac{1}{2}$ نماینده گروه iام

سپس میانگین از طریق فرمول (۲) محاسبه می شود.

مثال ۳ – جدول زیر توزیع فراوانی تعداد مراجعات ۵۰ نفر را به خانه بهداشت در طول یک سال نشان می دهد.

| حدود طبقات | $f_i$ |
|---|---|
| ۳–۵ | ۹ |
| ۶–۸ | ۱۰ |
| ۹–۱۱ | ۹ |
| ۱۲–۱۴ | ۱۰ |
| ۱۵–۱۷ | ۶ |
| ۱۸–۲۰ | ۶ |
| | ۵۰ |

می خواهیم متوسط(میانگین) مراجعات یک نفر را محاسبه کنیم. داریم:

| حدود طبقات | $f_i$ | $x_i$ | $f_i x_i$ |
|---|---|---|---|
| ۳–۵ | ۹ | ۴ | ۳۶ |
| ۶–۸ | ۱۰ | ۷ | ۷۰ |
| ۹–۱۱ | ۹ | ۱۰ | ۹۰ |
| ۱۲–۱۴ | ۱۰ | ۱۳ | ۱۳۰ |
| ۱۵–۱۷ | ۶ | ۱۶ | ۹۶ |
| ۱۸–۲۰ | ۶ | ۱۹ | ۱۱۴ |
| | ۵۰ | | ۵۳۶ |

$$\bar{x} = \frac{1}{50}\sum_{i=1}^{6} f_i x_i = \frac{536}{50} = 10.72$$

**خواص میانگین :**

۱) اگر متغیر X به صورت مجموع دو متغیر دیگر باشد یعنی داشته باشیم: $X_i = Y_i + Z_i$

در این صورت داریم:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n}(Y_i + Z_i) = \frac{1}{n}\sum_{i=1}^{n} Y_i + \frac{1}{n}\sum_{i=1}^{n} Z_i = \bar{Y} + \bar{Z} \quad \rightarrow \quad \bar{X} = \bar{Y} + \bar{Z}$$

۲ ) مجموع انحرافات داده ها از میانگین برابر صفر است زیرا :

$$\sum_{i=1}^{n}(X_i \quad \bar{X}) = \sum_{i=1}^{n} X_i \quad n\bar{X} = \sum_{i=1}^{n} X_i \quad n\frac{1}{n}\sum_{i=1}^{n} X_i = 0$$

۳ ) اگر به تمام داده ها مقدار ثابت a را اضافه  یا از تمام داده ها مقدار ثابت a را کم کنیم میانگین به اندازه a اضافه یا کم می شود زیرا داریم :

$$Y_i = X_i \pm a \rightarrow \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n}(X_i \pm a) = \frac{1}{n}\sum_{i=1}^{n} X_i \pm \frac{1}{n} na = \bar{X} \pm a$$

۴ ) هر گاه  تمام داده ها را در عدد ثابتی مانند b ضرب کنیم میانگین در b ضرب می شود زیرا :

$$Y_i = bX_i \rightarrow \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} bX_i = b\,\frac{1}{n}\sum_{i=1}^{n} X_i = b\bar{X}$$

به طور کلی :

$$Y_i = bX_i \pm a \rightarrow \bar{Y} = b\bar{X} \pm a$$

**ب) میانه**

اگر مشاهدات خود را به صورت غیر نزولی مرتب کنیم، میانه اندازه ای از صفت متغیر می باشد که در وسط قرار گرفته است و آن را با نماد $M_d$ می دهیم. اگر حجم نمونه فرد باشد و داده ها به صورت غیر نزولی مرتب شده باشد میانه داده ای است که در وسط قرار می گیرد. ولی اگر حجم نمونه زوج باشد میانه برابر با میانگین دو عدد وسطی است. به عبارت دیگر میانه برابر با عددی است که ۵۰ در صد داده ها که به صورت غیر نزولی مرتب شده اند، کمتر از آن و ۵۰ درصد داده ها بیشتر از آن باشد.

فرض کنید داده های $x_1,x_2,\dots,x_n$ را به صورت غیر نزولی مرتب کرده باشیم :

$$x_1 \le x_2 \le \dots \le x_k \le x_{k+1} \le \dots \le x_n$$

اگر n فرد و n=2k-1 باشد، در این صورت:

$$M_d = x_k \quad , \quad k = \frac{n+1}{2}$$

اگر n زوج و n=2k باشد در این صورت:

$$M_d = \frac{1}{2}(x_k + x_{k+1}),\ k = \frac{n}{2}$$

اگر داده های گسسته همراه با فراوانی داده شده باشد، برای محاسبه میانه ابتدا ستون مربوط به فراوانی تجمعی را تشکیل می دهیم و سپس با توجه به زوج یا فرد بودن حجم جامعه k را به صورت $\frac{n}{2}$ یا $\frac{n+1}{2}$ بدست می آوریم و مقدار k را با فراوانی های تجمعی مقایسه می کنیم .

**محاسبه میانه داده های طبقه بندی شده:** ابتدا ستون مربوط به فراوانی تجمعی تشکیل می دهیم پس عدد $k = \frac{n}{2}$ را با فراوانیهای تجمعی مقایسه می کنیم. اولین طبقه ای که فراوانی تجمعی آن بزرگتر یا مساوی k باشد را به عنوان طبقه میانه دار انتخاب می کنیم و میانه را از فرمول زیر محاسبه می کنیم:

$$M_d = L + \frac{\frac{n}{2} - F_c}{f_i} \times C$$

که در آن: $L$ کران پائین طبقه میانه دار
$F_c$ فراوانی تجمعی طبقه قبل از طبقه میانه دار
$f_i$ فراوانی مطلق طبقه میانه دار
$C$ فاصله طبقات می باشد.

مثال۸ – جدول توزیع فر اوانی وزن ۱۰۰ دانشجو در جدول زیر داده شده است میانه وزن دانشجویان را محاسبه کنید.

| حدود طبقات | $f_i$ | $F_c$ |
|---|---|---|
| ۶۰–۶۲ | ۵ | ۵ |
| ۶۳–۶۵ | ۱۸ | ۲۳ |
| ۶۶–۶۸ | ۴۲ | ۶۵ |
| ۶۹–۷۱ | ۲۷ | ۹۲ |
| ۷۲–۷۴ | ۸ | ۱۰۰ |
| | ۱۰۰ | |

حل : عدد $k=\frac{n}{2}=\frac{100}{2}=50$ را با ستون فراوان تجمعی مقایسه می کنیم. اولیه طبقه ای که فراوانی تجمعی آن بزرگتر یا مساوی ۵۰ است، طبقه سوم می باشد که به عنوان طبقه میانه دار انتخاب می شود. حال با توجه به فرمول محاسبه میانه داریم:

$$\frac{n}{2}=50\,,L=65.5\,,F_c=23\,,f_i=42\,,C=3$$
$$M_d=65.5+\frac{50-23}{42}\times 3=67.4$$

## ج ) مد یا نما :

اندازه ای از صفت متغیر است که فراوانی آن ماکسیمم یا بیشترین است و آن را با نماد $M_o$ نشان می دهیم. اگر صفت یک صفت کیفی با توزیع فراوانی زیر باشد:

| حالات | $E_1$ | $E_2$ | … | $E_k$ | … | $E_n$ |
|---|---|---|---|---|---|---|
| فراوانی | $f_1$ | $f_2$ | … | $f_k$ | … | $f_n$ |

فرض کنید $f_k$ ماکسیمم $f_i$ها باشد. در این صورت: $M_o=E_k$

**توجه** : اگر پیش از یک نما در مشاهدات خود داشته باشیم، توزیع را چند نمایی می گوئیم و در این صورت مد مرکز تمرکز نیست.

محاسبه مد در داده های طبقه بندی شده: ابتدا طبقه ای که فراوانی آن ماکسیمم است را مشخص می کنیم و آن را طبقه مددار می نامیم و مد را از فرمول زیر حساب می کنیم:

$$M_o=L+\frac{d_1}{d_1+d_2}\times C$$

که در آن : $L$ کران پائین طبقه مدار
$d_1$ تفاضل فراوانی طبقه مد دار از فراوانی طبقه قبل
$d_2$ تفاضل فراوانی طبقه مد دار از فراوانی طبقه بعد
$C$ طول فاصله طبقه مد دار

مثال ۹ – برای مثال قبل مد را محاسبه کنید :

حل : چون طبقه سوم بیشترین فراوانی را داراست لذا طبقه سوم طبقه مد دار می باشد و با توجه به فرمول داریم:

$$d_1=42-18=24\,,d_2=42-27=15\,,L=65.5\,,C=3$$

$$M_o = 65.5 + \frac{24}{24+15} \times 3 = 67.35$$

رابطه تجربی بین میانگین میانه و مد : در توزیع هایی که تقریباً متقارن باشد رابطه تجربی که رابطه پیرسن معروف است برقرار می باشد: $\bar{x} - M_o = 3(\bar{x} - M_d)$

**توجه** : رابطه فوق را برای صحت محاسبات بکار می بریم و هیچ گاه برای محاسبه یکی از شاخصهای مرکزی بر حسب دوتای دیگر از این فرمول استفاده نمی کنیم .

## ۲ – شاخصهای پراکندگی :

توزیع های فراوانی ناشی از پراکندگی صفات می باشد و میزان پراکندگی بوسیله شاخصهای مرکزی ظاهر نمی شوند . بنابراین از شاخصهای پراکندگی برای بررسی میزان پراکندگی یک توزیع فراوانی استفاده می کنیم .

**الف) دامنه**

دامنه یا میدان تغییرات صفت که آن را با نماد R نشان می دهیم عبارت است از تفاضل بزرگترین داده از کوچکترین داده :
کوچکترین داده – بزرگترین داده = R

حسن دامنه در آن است که محاسبه آن ساده است و عیب آن این است که دامنه به عنوان یک پارامتر پراکندگی فقط به دو مقدار بزرگ و کوچک بستگی دارد و بقیه داده ها دخالتی در مقدار R ندارند.

**ب) واریانس و انحراف معیار**

یکی از مهمترین پارامترها درعلم آمار واریانس است و عبارت است از میانگین مربع انحرافات از میانگین و آن را با $S^2$ نشان می دهیم. واریانس جامعه را با نماد $\sigma^2$ نشان می دهیم و آن را با استفاده از فرمولهای زیر محاسبه می کنیم :

الف ) داده های طبقه بندی نشده :

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

ب) داده های طبقه بندی شده :

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

مثال ۱۱ : جدول توزیع فراوانی نمرات درس آمار ۱۰۰ دانش آموز بهورز به صورت زیر مضروض است. مطلوب است محاسبه واریانس:

| حدود طبقات | $f_i$ |
|---|---|
| ۶۰–۶۲ | ۵ |
| ۶۳–۶۵ | ۱۸ |
| ۶۶–۶۸ | ۴۲ |
| ۶۹–۷۱ | ۲۷ |
| ۷۲–۷۴ | ۸ |
| | ۱۰۰ |

حل:

| حدود طبقات | $f_i$ | $x_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ | $f_i(x_i - \bar{x})^2$ | $f_i x_i$ |
|---|---|---|---|---|---|---|
| ۶۰–۶۲ | ۵ | ۶۱ | −۶/۴۵ | ۴۱/۶۰۲۵ | ۲۰۸/۰۱۲۵ | ۳۰۵ |
| ۶۳–۶۵ | ۱۸ | ۶۴ | −۳/۴۵ | ۱۱/۹۰۲۵ | ۲۱۴/۲۴۵۰ | ۱۱۵۲ |
| ۶۶–۶۸ | ۴۲ | ۶۷ | ۰/۴۵ | ۰/۲۰۲۵ | ۸/۵۰۵۰ | ۲۸۱۴ |
| ۶۹–۷۱ | ۲۷ | ۷۰ | ۲/۵۵ | ۶/۵۰۲۵ | ۱۷۵/۵۶۷۵ | ۱۸۹۰ |
| ۷۲–۷۴ | ۸ | ۷۳ | ۵/۵۵ | ۳۰/۸۰۲۵ | ۲۶۴/۴۲ | ۵۸۴ |
| | ۱۰۰ | | | | ۸۵۲/۷۵ | ۶۷۴۵ |

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i = \frac{6745}{100} = 67.45$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{k} f_i (x_i - \bar{x})^2 = \frac{852.45}{100} = 8.5245$$

**انحراف معیار یا انحراف استاندارد:** عبارت است از جذر مثبت واریانس و آن را با نماد $S$ نشان می دهیم.

مثال ۱۲– در مثال فوق انحراف معیار را محاسبه کنید.

حل :

$$S = \sqrt{8.5275} = 2.92$$

**خواص واریانس**

۱) اگر به تمام داده ها مقدار ثابت $b$ را اضافه کنیم و اریانس تغییر نمی کند . به عبارت دیگر $X$ و $Y$ دارای واریانس برابر می باشند اگر : $Y=X+b$

۲) اگر تمام داده ها را در عدد ثابتی مانند $a$ ضرب کنیم، واریانس جدید در $a^2$ ضرب می شود. زیرا اگر $Y=aX$ آنگاه:

$$V(Y) = V(aX) = \frac{1}{n}\sum_{i=1}^{n}(aX_i - a\bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} a^2(X_i - \bar{X})^2 = a^2\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = a^2 V(X)$$

نکته: به سادگی می توان نشان داده که:

$$V(aX + b) = a^2 V(X)$$

توجه : فرمول زیر محاسبه واریانس را ساده تر می کند: (بخصوص در مواردی که $\bar{x}$ اعشاری باشد)

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

**ج) ضریب تغییرات یا ضریب واریانس یا پراکندگی**

با نماد $C.V$ نشان داده می شود و عبارت است از :

$$C.V = \frac{S}{\bar{x}} \times 100$$

که به صورت درصد بیان می شود و برای مقایسه در جمعیت استفاده می شود.

# بخش سوم- احتمالات

## ۱- تعاریف و کلیات

به طور کلی اصطلاح احتمال مربوط به وقوع پیش آمدهای تصادفی است و یا به عبارتی احتمالات عبارت از بررسی آزمونهای تصادفی است و به طور خلاصه در بررسی مسائل احتمال با آزمایشها و پیش آمدها سر وکار داریم. در مباحث قبل در مورد آمار توصیفی بحث شد و پارامترها یا شاخصهایی را براساس نمونه گرفته شده از جامعه بدست آوردیم. حال اگر بخواهیم نتایج بدست آمده از نمونه را به جامعه تعمیم دهیم و یا به عبارتی بخواهیم از پارامتر های بدست آمده در نمونه، برآوردی درباره پارامترهای جامعه داشته باشیم، این تعمیمات استنباطی براساس احتمال یا درستنمایی پیش آمدها متکی می باشد. به عبارت دیگر فرق احتمال و آمار در این است که در آمار از کل به جزء می رسیم، ولی دراحتمال جزء را به کل تعمیم می دهیم . پیش از هرچیز به بیان چند تعریف می پردازیم :

**آزمایش :** عبارت است از هر روشی که امکان جمع آوری داده ها باشد.

**آزمایش تصادفی :** آزمایشی است که نتایج آن از قبل به طور قطع مشخص نباشد.

**فضای نمونه :** مجموعه تمام نتایج ممکن در یک آزمایش را فضای نمونه می نامیم و آن را با حرف S نمایش می دهیم .

مثال۱ – تاسی را پرتاب می کنیم فضای نمونه آن را مشخص کنید .

حل : $S=\{1,2,3,4,5,6\}$

مثال۲ – در پرتاب ۲ سکه فضای نمونه را مشخص کنید :

حل : اگر H نشان دهنده رو و T نشان دهنده پشت سکه باشد، فضای نمونه بصورت زیر است :

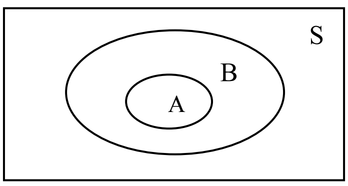
$S=\{HH,HT,TH,TT\}$

**پیشامد :** هر مجموعه از نتایج یک آزمایش را پیشامد می گوئیم به عبارت دیگر پیشامد زیر مجموعه ای از فضای نمونه ای است .

پیشامد $\emptyset$ را پیشامد غیر ممکن و پیشامد S را پیشامد حتمی می نامیم. پیشامدهای یک عضوی را پیشامد ساده و پیشامدهایی که شامل بیش از یک عضو می باشد راپیشامد مرکب می نامیم.

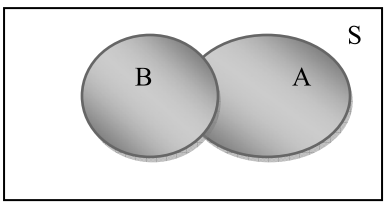
## ۲- نمودار ون و جبر پیشامدها

فرض کنید AوB دو پیشامد از فضای نمونه S باشد:

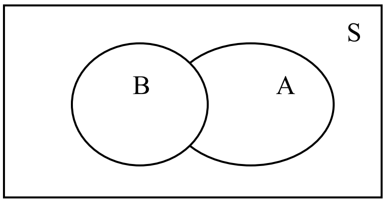



۱- A را زیر مجموعه B می نامیم و آن را با $A \subseteq B$ نشان می دهیم هرگاه هر عضو A عضوی از B باشد بنابر این اگر $A \subseteq B$ ، آنگاه رخ دادن A ، لزوما رخ دادن B را نتیجه می دهد.

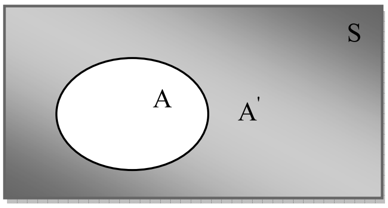
۲ – اگر $A \subseteq B$ و $B \subseteq A$ آنگاه A وB را دوپیشامد یکسان می نامیم.



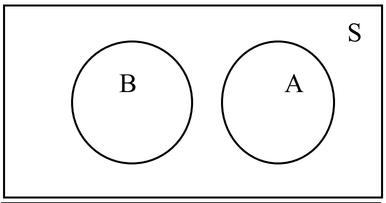

۳ –اجتماع دو پیشامد AوB را با نماد $A \cup B$ نشان می دهیم و عبارت است از پیشامد هایی که در اثر وقوع پیشامد A یا پیشامد B یا هردو رخ می دهد. به عبارت دیگر پیشامد$A \cup B$ فقط و فقط وقتی رخ می دهد که پیشامد A و یا B و یا هردو رخ دهد.



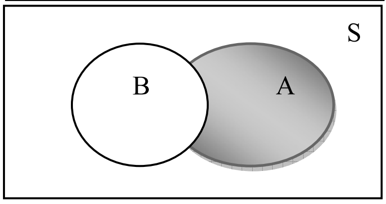

۴ – اشتراک دو پیشامد AوB را با $A \cap B$ نشان می دهیم و عبارت است از پیشامدی که در اثر وقوع پیشامد های AوB رخ دهد. به عبارت دیگر پیشامد $A \cap B$ فقط و فقط وقتی رخ می دهد که پیشامدهای AوB با هم رخ دهند.




۵ – پیشامد $A'$ متمم پیشامد A نامیده می شود اگر پیشامد $A'$ فقط و فقط وقتی رخ دهد که پیشامد A رخ ندهد.




۶ – AوB را دو پیشامد ناسازگار می نامیم اگر وقوع توام آنها غیر ممکن باشد با عبارت دیگر $A \cap B = \emptyset$




۷ – تفاضل دو پیشامد AوB پیشامدی است که در اثر وقوع پیشامد A و عدم وقوع پیشامد B رخ دهد و آن را با A-B نشان می دهیم به عبارت دیگر: $A - B = A \cap B'$

قواعدی برای پیشامدها :

اگر AوBوC پیشامدهایی از فضای نمونه S باشد قوانین زیر در مورد آنها برقرار است:

$A \subseteq A$

$\emptyset \subseteq A$

$A \subseteq S$

$A \subseteq B , B \subseteq C \rightarrow A \subseteq C$

$A = B , B = C \rightarrow A = C$

$A = B \rightarrow A' = B'$

$S' = \emptyset$

$\emptyset' = S$

$A \cup A = A$

$A \subseteq B \rightarrow A \cup B = B$

$A \subseteq B \rightarrow A \cap B = A$

$A \cup A' = S$

$(A \cup B) \cup C = A \cup (B \cup C)$

$(A \cap B) \cap C = A \cap (B \cap C)$

$A \cap B = B \cap A$

$A \cup B = B \cup A$

$A \cap A' = \emptyset$

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

$(A \cap B)' = A' \cup B'$

$(A \cup B)' = A' \cap B'$

## ۳- احتمال

تعریف احتمال : مجموع وزن های تمام عضوهای موجود در پیشامد A را احتمال پیشامد A می نامیم آن را با P(A) نشان می دهیم. با توجه به تعریف پیشامد A داریم:

۱- $P(\emptyset)=0$ ۲ - $P(S)=1$ ۳ - $0 \le P(A) \le 1$

تعریف : پیشامد های $E_1, E_2, \dots, E_n$ راپیشامد های ناسازگار گوئیم هرگاه وقوع همزمانی هر دو پیشامد ، غیر ممکن باشد. به عبارت دیگر: $E_i \cap E_j = \emptyset \quad , i \neq j , 1 \le i, j \le n$

در نتیجه: $P(E_i \cap E_j) = 0$

احتمال کلاسیک : اگر دو گروه کامل پیشامد های ناسازگار پیشامدها هم شانس باشند و در این صورت احتمال پیشامدی مانند A را که با P(A) نشان می دهیم عبارت است از: $P(A) = \frac{n(A)}{n(S)} = \frac{\text{تعداد حالات مساعد}}{\text{تعداد حالات ممکن}}$

توجه : در تعریف فوق که احتمال بدون انجام آزمایش محاسبه می شود باید سه شرط زیر برقرار باشد:

۱ - وقوع یکی از پیشامدها حتمی باشد.

۲ - پیشامدها ناسازگار باشد.

۳ - پیشامدها همتراز یا هم شانس باشد.

مثال : یک تاس سالم را یک بار پرتاب می کنیم مطلوب است محاسبه احتمال این که :

الف: ۵ ظاهر شود. ب : زوج ظاهر شود.

حل : الف ) اگر پیشامد A وقوع عدد ۵ باشد: A={۵} و n(A)=۱ و چون S={۱,۲,۳,۴,۵,۶} و n(S)=۶

پس : $P(A) = \frac{n(A)}{n(S)} = \frac{1}{6}$

ب) اگر پیشامد B وقوع عدد زوج باشد: B={۲,۴,۶} و n(B)=۳

پس : $P(B) = \frac{n(B)}{n(S)} = \frac{1}{2}$

اصول موضوع احتمالات :

۱ - اگر $A'$متمم A باشد آنگاه: $P(A') = 1 - P(A)$

۲ - $P(S) = 1$

۳ - اگر Aو B دو پیشامد ناسازگار باشند : $P(A \cup B) = P(A) + P(B)$

۴ - $A \subseteq B \rightarrow P(A) \le P(B)$

۵ - $A = B \rightarrow P(A) = P(B)$

۶ - اگر Aو B دو پیشامد دلخواه باشد آنگاه : الف) $P(A - B) = P(A) - P(A \cap B)$

ب) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

مثال ۴ : یک گروه ۳۰ نفری متشکل از ۱۰ مرد و ۲۰زن ، رنگ چشم نصف مردان و نصف زنان مشکی باشد یک نفر رابطور تصادفی انتخاب می کنیم مطلوب است احتمال آنکه یک مرد یا یک چشم مشکی انتخاب شود.

حل : فرض کنید A پیشامد مرد بودن و B پیشامد مشکی بودن رنگ چشم باشد :

$P(A) = \frac{10}{30} = \frac{1}{3} \qquad P(B) = \frac{15}{30} = \frac{1}{2} \qquad P(A \cap B) = \frac{5}{30} = \frac{1}{6}$

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{3} + \frac{1}{2} - \frac{1}{6} = \frac{2}{3}$

# بخش چهارم – شاخص های بهداشتی

## ۱ – مقدمه

شاخص های بهداشتی اطلاعات خلاصه شده ای هستند که به منظور پاسخگویی به سوالاتی در زمینه برنامه ریزی و مدیریت برنامه های بهداشتی جمع آوری می شوند . شاخص های بهداشتی برای ارزیابی وضعیت جمعیت پایش برنامه ها و نتایج اجرایی آنها و ارزیابی اثر بخشی و نتایج نهایی یک برنامه بکار می رود . بسیاری از شاخص های بهداشت مسائلی را نمایان می کنند که نیاز به مداخله دارند تا اینکه یک ابزار تشخیص باشند.

یک شاخص بهداشتی می تواند بیانگر مسائلی از این قبیل باشد:

الف) وقوع یک واقعه : نظیر یک تولد، یک مرگ یا یک زایمان خطرناک

ب) شیوع یک خصوصیت رفتاری در یک فرد: مانند استفاده از یک روش پیشگیری بوسیله یک زن با وزن کم بدو تولد در یک کودک

ج) ویژگی یک مرکز بهداشتی درمانی : نظیر مرکز بهداشتی درمانی که مراقبتهای دوران بارداری را انجام می دهند.

به هر حال شاخص های بهداشتی کار برد فراوان دارند از جمله :

۱) پایش تغییر در زمان ( روند زمانی ) نظیر تغییر در نسبت افرادی که مبتلا به بیماریهای آمیزشی بوده و به مرکز بهداشتی درمانی مراجعه و مورد ارزیابی و آموزش قرار گرفته اند .

۲) پایش اختلاف بین گروههایی از جمعیت مثلاً نسبت زایمانهای انجام شده توسط ماماهای آموزش دیده در زنان با سطح تحصیلات متفاوت.

۳) پایش میزان دسترسی به اهداف مانند درصد مادران بارداری که قرص آهن، فولیک اسید را دریافت کردند نسبت به هدفی که داشته ایم.

۴) پایش اختلاف بین مراکز بهداشتی درمانی در مناطق جغرافیایی مختلف بعنوان نمونه پوشش مراقبتهای کودکان در مراکز بهداشتی درمانی مختلف.

شاخص ها معمولا بصورت کمی و کسری بیان می شوند ولی می توان آنها را بصورت کیفی یا در قالب اعداد مطلق نیز بیان نمود. برای مثال هنگامی که بررسی تغییرات مصرف سیگار در یک جامعه با جمعیت نسبتاً ثابت مورد نظر است می توان از عدد مطلق میزان سیگار به فروش رفته بعنوان شاخص مصرف سیگار اسفتاده کرد. وجود یا عدم وجود قانون برای گزارش اجاری بیماریها در حکم یک شاخص کیفی برابر ارزیابی نظام مراقبت بیماریهاست .

## ۲ – میزان ها و نسبت ها :

به طور کلی در تعیین شاخصهای بهداشتی از دو کلمه میزان و نسبت استفاده می کنیم. به عنوان مثال، میزان مرگ و میر مادران باردار یا نسبت گروههای سنی در جمعیت تحت پوشش یک واحد خاص. برای تفسیر این دو اصطلاح می توان گفت در شاخصهایی که در آنها نسبت بکار می رود عدد شاخص در عدد ثابتی ضرب نمی شود. مثلاً اگر بخواهیم شاخص نسبت جنسی در بدو تولد را در یک منطقه خاص محاسبه کنیم تعداد متولدین پسر در یک مقطع زمانی ( یک سال ) تقسیم بر تعداد

متولدین دختر می شود . این عدد معمولاً حول عدد یک در گردش است. اگر تعداد متولدین پسر بیشتر از دختر باشد شاخص بدست آمده بزرگتر یک است و بالعکس. اصطلاح نسبت معمولاً برای شاخصهایی بکار می رود، که از لحاظ عددی صورت و مخرج نزدیک هم هستند و حاصل عدد معنی داری می باشد. ولی اگر صورت نسبت به مخرج عدد کوچکی باشد که با تقسیم صورت به مخرج عدد بدست آمده بسیار کوچک و نزدیک به صفر باشد، در این صورت برای محسوس بودن شاخص عدد بدست آمده را در عددی ثابت ضرب می کنیم. این عدد بستگی به تعریف شاخص دارد و از ۱۰۰ به بالا می باشد. حال اگر آن عدد ۱۰۰ باشد، شاخص مورد نظر اصطلاح در صد را به خود اختصاص می دهد. به طور مثال برای بدست آوردن شاخص درصد افراد زیر ۵ سال در یک منطقه خاص تعداد افراد زیر ۵ سال منطقه را بر کل جمعیت تحت پوشش آن منطقه تقسیم می کنیم. حاصل را در عدد ۱۰۰ ضرب می کنیم اما برای بدست آوردن شاخص میزان مرگ کودکان زیر ۵ سال آن منطقه حاصل تقسیم تعداد مرگ کودکان زیر ۵ سال بر جمعیت زیر ۵ سال را در عدد ۱۰۰۰ ضرب می کنیم استدلال این امر این است که، چون تعداد مرگ و میر کودکان در حال حاضر نسبت به جمعیت عدد بسیار کمی می باشد، از طرفی مرگ و میر کودکان شاخص نسبتاً مهمی برای کارشناسان مربوطه می باشد، با ضرب عدد بدست آمده در عدد ۱۰۰۰ به طور محسوسی مشخص می شود .

## ۳ – نحوه بدست آوردن شاخص های مهم بهداشتی

**الف ) شاخص های جمعیتی**

۱- درصد گروه سنی زیر یکسال: $\dfrac{\text{۱۰۰× جمعیت زیر یکسال}}{\text{کل جمعیت}}$

۲- درصد گروه سنی زیر ۵ سال: $\dfrac{\text{۱۰۰× جمعیت زیر ۵ سال}}{\text{کل جمعیت}}$

۳- درصد گروه سنی زیر ۱۵سال: $\dfrac{\text{۱۰۰× جمعیت زیر ۱۵سال}}{\text{کل جمعیت}}$

۴- درصد گروه سنی ۱۵ تا ۶۴ سال: $\dfrac{\text{۱۰۰× جمعیت ۱۵ تا۶۴ سال}}{\text{کل جمعیت}}$

۵- درصد گروه سنی ۶۵ و بالاتر سال: $\dfrac{\text{۱۰۰× جمعیت ۶۵ سال بالاتر}}{\text{کل جمعیت}}$

۶- درصد سرباری: $\dfrac{\text{۱۰۰× جمعیت زیر ۱۵ سال + جمعیت ۶۵ سال و بالاتر سال}}{\text{جمعیت ۱۵تا ۶۴ سال}}$

۷- درصد زنان ۱۰-۴۹ سال شوهر دار: $\dfrac{\text{۱۰۰× جمعیت زنان ۱۰-۴۹ سال شوهردار}}{\text{کل جمعیت زنان ۱۰-۴۹ سال}}$

۸- میزان خان تولد (به طور معمول در هزار): ۱۰۰× تعداد تولد های زنده

۹- میزان باروری عمومی (به طور معمول در هزار)

$$\frac{\text{تعداد موالید زنده} \times ۱۰۰}{\text{جمعیت زنان ۱۰تا۴۹سال}}$$

کل جمعیت

۱۰-میزان باروری اختصاصی سنی (بطور معمول در هزار برای آگاهی از میزان حاملگی در هر گروه سنی معین) .

$$\frac{\text{تعداد موالید زنده در هر گروه سنی زنان} \times ۱۰۰}{\text{جمعیت زنان آن گروه سنی}}$$

۱۱-میزان باروری کلی (به طور معمول در هزار

$$\frac{\text{مجموع میزانهای باروری اختصاصی سنی} \times ۵}{۱۰۰۰}$$

۱۲-نسبت جنس در بدو تولد

$$\frac{\text{تعداد متولدین زند پسر} \times ۱۰۰}{\text{تعداد متولدین زنده دختر}}$$

۱۳-میزان خام مرگ (به طور معمول در هزار )

$$\frac{\text{تعداد مرگ و میر} \times ۱۰۰}{\text{کل جمعیت}}$$

۱۴-میزان مرگ نوزادان زیر یکماه (به طور معمول در هزار تولد )

$$\frac{\text{تعداد مرگ نوزادان زیر یکماه} \times ۱۰۰}{\text{موالید زنده}}$$

۱۵-میزان مرگ نوزادان زیر یکسال (به طور معمول در هزار تولد )

$$\frac{\text{تعداد مرگ نوزادان زیر یکسال} \times ۱۰۰}{\text{موالید زنده}}$$

۱۶- میزان مرگ کودکان زیر ۵سال: این شاخص به دو روش محاسبه می شود:

الف ) بر حسب جمعیت کودکان زیر ۵ سال

$$\frac{\text{تعداد مرگ کودکان زیر ۵ سال} \times ۱۰۰۰}{\text{جمعیت زیر ۵ سال}}$$

ب ) بر حسب موالید زنده

$$\frac{\text{تعداد مرگ کودکان زیر ۵ سال} \times ۱۰۰۰}{\text{موالید زنده}}$$

۱۷-میزان اختصاصی مرگ مودکان زیر ۵ سال بر حسب علت فوت (به طور معمول در ۱۰۰۰ یا ۱۰۰۰۰)

$$\frac{\text{تعداد مرگ کودکان زیر ۵ سال به علت موردنظر} \times ۱۰۰۰}{\text{جمعیت زیر ۵ سال}}$$

۱۸- رشد طبیعی جمعیت :

میزان مرگ خام – میزان خام تولد = رشد طبیعی جمعیت

**ب) شاخص های بهداشت خانواده**

۱- درصد افراد تحت پوشش تنظیم خانواده (کل وسایل )

$$\frac{\text{تعداد تحت پوشش تنظیم خانواده (کل وسایل)} \times ۱۰۰}{\text{جمعیت زنان ۱۰تا۴۹ سال همسردار}}$$

۲- درصد افراد تحت پوشش تنظیم خانواده (روشهای مدرن )

$$\frac{\text{تعداد تحت پوشش تنظیم خانواده (روشهای مدرن)} \times 100}{\text{جمعیت زنان ۱۰تا۴۹ سال همسردار}}$$

۳- درصد افراد استفاده کننده از وسیله خاص تنظیم خانواده

$$\frac{\text{تعداد استفاده کننده از وسیله مورد نظر} \times 100}{\text{جمعیت تحت پوشش تنظیم خانواده (روشهای مدرن)}}$$

۴- درصد زایمانهای انجام شده در منزل یا بین راه

$$\frac{\text{تعداد زایمانهای انجام شده در منزل یا بین راه} \times 100}{\text{تعداد کل زایمانها}}$$

۵- درصد زایمانهای انجام شده در بیمارستان

$$\frac{\text{تعداد زایمانهای انجام شده در بیمارستان} \times 100}{\text{تعداد کل زایمانها}}$$

۶- درصد زایمانهای انجام شده در منزل توسط مامای دوره ندیده

$$\frac{\text{تعداد زایمانهای انجام شده در منزل توسط ماما دوره ندیده} \times 100}{\text{تعداد کل زایمانها در منزل}}$$

۷- درصد زایمان به روش طبیعی

$$\frac{\text{تعداد زایمانهای طبیعی} \times 100}{\text{تعداد کل زایمانها}}$$

۸- درصد زایمان به روش سزارین

$$\frac{\text{تعداد زایمانهای به روش سزارین} \times 100}{\text{تعداد کل زایمانها}}$$

۹- درصد زنانی که مراقبت های پس از زایمان را دریافت کرده اند

$$\frac{\text{تعداد که مراقبت های پس از زایمان را دریافت کرده اند} \times 100}{\text{تعداد کل زایمانها}}$$

۱۰- درصد زنانی که بارداری که قرص آهن / مولتی ویتامین /اسید فولیک را دریافت کرده اند

$$\frac{\text{تعداد که بارداری که قرص آهن / مولتی ویتامین /اسید فولیک را دریافت کرده اند} \times 100}{\text{تعداد کل زنان باردار مراقبت شده}}$$

۱۱- در صد زایمانهای انجام شده توسط مامای دوره دیده (ماما روستا)

$$\frac{\text{تعداد زایمانهای انجام شده توسط مامای دوره دیده (ماما روستا)} \times 100}{\text{تعداد کل زایمانها}}$$

۱۲- میزان مرگ مادران بدلیل عوارض بارداری و زایمان(به طور معمول در صدهزار)

$$\frac{\text{۱۰۰۰۰۰} \times \text{تعداد مرگ مادران دلیل عوارض بارداری و زایمان}}{\text{متولدین زنده}}$$

۱۳- در صد خانوارهای استفاده کننده از نمک یددار

$$\frac{\text{۱۰۰} \times \text{تعداد خانوارهای استفاده کننده از نمک یددار}}{\text{تعداد کل خانوار}}$$

۱۴- میزان کودکان زیر ۶سال دچار اختلال رشد (به طور معمول در هزار)

$$\frac{\text{۱۰۰۰} \times \text{تعداد کودکان زیر ۶سال دچار اختلال رشد (به طور معمول در هزار}}{\text{جمعیت کودکان زیر ۶ سال}}$$

۱۵- میزان کودکان زیر ۶سال دچار سوء تغذیه(به طور معمول در هزار)

$$\frac{\text{۱۰۰۰} \times \text{تعداد کودکان زیر زیر ۶سال دچار سوء تغذیه(به طور معمول در هزار)}}{\text{جمعیت کودکان زیر ۶ سال}}$$

**ج) شاخص های بهداشت محیط و حرفه ای**

۱- در صد خانواده های تحت پوشش استفاده کننده از توالت های بهداشتی)

$$\frac{\text{۱۰۰} \times \text{تعداد خانواده های تحت پوشش استفاده کننده از توالت های بهداشتی)}}{\text{تعداد کل خانواده های تحت پوشش}}$$

۲- در صد استفاده از روشهای بهداشتی دفع زباله در روستا

$$\frac{\text{۱۰۰} \times \text{تعداد خانواده های تحت پوشش خانه ای بهداشت که دفع زباله آنها به روش بهداشتی صورت می گیرد}}{\text{تعداد کل خانواده های تحت پوشش}}$$

۳- در صد استفاده کننده از روشهای بهداشتی دفع فضولات حیوانی در روستا

$$\frac{\text{۱۰۰} \times \text{تعداد خانواده های تحت پوشش خانه ای بهداشت که دفع فضولات حیوانی آنها به روش بهداشتی صورت می گیرد}}{\text{تعداد کل خانواده های تحت پوشش}}$$

۴- در صد کارکنان مراکز و مکانهای عمومی و دارای کارت بهداشتی

$$\frac{\text{۱۰۰} \times \text{تعداد کارکنان مکانهای عمومی و دارای کارت بهداشتی}}{\text{تعداد کل کارکنان مکانهای عمومی تحت پوشش}}$$

۵- در صد کارگاههای تحت پوشش (به تفکیک خانگی و غیر خانگی )

۱۰۰× تعداد کارگاههای تحت پوشش
تعداد کار گاهای شناسایی شده

**د) شاخص های بهداشت مدارس**

۱- در صد دانش آموزان برخوردار از تسهیلات بهداشتی کافی در مدارس

۱۰۰× تعداد دانش آموزانی که در مدرسه به تسیلات بهداشتی درمانی دسترسی دارند
کل دانش آموزان تحت پوشش

توجه : تسهیلات کافی بهداشتی : یک توالت بهداشتی برای هر ۴۵ دانش آموز، یک دستشویی برای هر ۴۵ دانش آموز، یک آبخوری برای هر ۷۵ دانش آموز و آب سالم.

۲- در صد دانش آموزان برخوردار از بوفه بهداشتی در مدارس

۱۰۰× تعداد دانش آموزانی که در مدرسه دارای بوفه بهداشتی
کل دانش آموزان تحت پوشش

**ه) شاخص های مبارزه با بیماریها**

۱- درصد پوشش واکسیناسیون نوبت اول واکسن ب ث ژ / نوبت دوم سرخک / نوبت سوم ثلاث / نوبت سوم پولیو / نوبت سوم هپاتیت در کودکان زیر یکسال

۱۰۰× تعداد واکسیناسیون انجام شده (به تفکیک نوع واکسن ونوبت )
جمعیت کودکان زیر یکسال

۳- درصد پوشش واکسیناسیون کزاز زنان باردار

۱۰۰× تعداد زنان بارداری که ایمن سازی دوران کودکی آنها کامل است یا در هنگام جاملگی به آنها دوبارواکسن کزاز تزریق شده است
جمعیت کودکان زیر یکسال یا جمعیت زنان باردار

∞∞∞ **پایان** ∞∞∞

منابع:

۱- آمار و احتمالات تالیف دکتر مسعود نیکوکار
۲- اینترنت
۳- شاخص های بهداشتی گرد آوری شده به وسیله دکتر سیروس پیله وری

# فهرست



انتظار می رود فراگیر پس از مطالعه این جزوه :

۱- مفاهیم اولیه آماری و نحوه جمع آوری داده ها را بیان کند.
۲- جداول توزیع فراوانی را به تفکیک نوع داده ها رسم نماید.
۳- نمودارهای آماری (سوزنی- میله ای- ستونی- چندضلعی) را برای هر نوع داده رسم کند.
۴- شاخص های مرکزی (میانگین- میانه- مد یا نما) را محاسبه نماید و کاربرد هریک را بیان کند.
۵- شاخص های پراکندگی داده ها (دامنه - واریانس- انحراف معیار) را محاسبه و در مسائل مختلف استفاده نماید.
۶- نمودارهای ون را برای جبر (عملیات ریاضی ) پیشامدها رسم و آنها را تفسیر کند.
۷- مفهوم احتمال را بیان و قوانین مربوط به آن را یاد بگیرد.
۸- مفهوم میزان و نسبت و درصد و کاربرد آنها را در شاخص های بهداشتی بداند.
۹- نحوه محاسبه شاخص های بهداشتی مربوطه را به طور کامل یاد بگیرد.

بسم الله الرحمن الرحيم