

اگرچه نیت خوبی است زیستن ...  
اما خوشا که دست به تصمیم بهتری بزنیم!

 [www.konkursara.com](http://www.konkursara.com)

 ۰۲۱۵۵۷۵۶۵۰۰

دانلود بهترین جزوات در

**کنکورسرا**

کنکورسرا

مرجع تخصصی قبولی آزمون فرهنگیان و آزمون استخدامی آموزش و پرورش

# آمار و احتمال

پایه می یازدهم « رشته می ریاضی فنریک »

فصل ۴ : آمار استنباطی

## درس ۱: گردآوری اطلاعات آماری

بدیهی است که انجام هر کار آماری برای شناخت حداقل یکی از ویژگی های یک مجموعه از افراد یا اشیاء است. اولین اقدام در چنین فعالیت هایی، گردآوری اطلاعات (داده ها) پیرامون جامعه آماری (یا نمونه) می باشد. در فرآیند گردآوری اطلاعات به اصطلاحاتی از قبیل اصطلاحات زیر روبرو می شویم. تعریف این اصطلاحات به صورت زیر است:

**داده:** داده ها واقعیت هایی درباره ی یک موضوع هستند که در محاسبه، استنباط یا برنامه ریزی به کار می روند.

مثال: وزن هر یک از نوزادانی که در ماه گذشته در شهر خاصی به دنیا آمده اند.

**واحد آماری:** به هر یک از افراد یا اشیاء گویند که داده های مربوط به آنها در یک بررسی آماری گردآوری می شوند.

مثال: هر کدام از نوزادان که در ماه گذشته در شهر خاصی به دنیا آمده اند.

مثال: وقتی گفته می شود که معدل علی محمدی  $14/5$  است. شخص علی محمدی واحد آماری و عدد  $14/5$  یک داده است.

مثال: وقتی گفته می شود، تعداد دبیران ریاضی شهرستان رامهرمز ۳۴ نفر است. شهرستان رامهرمز واحد آماری و عدد ۳۴ داده است.

مثال: وقتی گفته می شود که محسن محمدی از خدمات رستوران گلها ناراضی است. شخص محسن محمدی واحد آماری و واژه ی ناراضی داده است.

**جامعه آماری:** مجموعه ی کل واحدهای آماری را جامعه آماری می نامند.

مثال: مجموعه ی کل نوزادانی که در ماه گذشته در شهر خاصی به دنیا آمده اند.

**نمونه:** هر زیر مجموعه از جامعه آماری را که با روش مشخصی انتخاب شده باشند، یک نمونه آماری می نامند.

مثال: نوزادانی که روز های زوج ماه گذشته به دنیا آمده اند.

**نمونه گیری:** فرآیند انتخاب نمونه ای از یک جامعه آماری، با منظور تعمیم اطلاعات آن به جامعه است.

مثال : فرآیند انتخاب نوزادانی که روز های زوج ماه گذشته به دنیا آمده اند.

از آنجا که جوامع آماری معمولاً از نظر حجم و وسعت جغرافیایی زیادی برخوردارند و محققان نمی توانند به تمام آنها مراجعه کنند، بنابراین ناگزیرند به انتخاب جمعی از آنها به عنوان نمونه و تعمیم نتایج آن به جامعه‌ی مورد مطالعه اکتفا کنند. مثلاً محقق نمی تواند تمام روستاهای یک استان یا حتی یک شهرستان را مطالعه کند، یا برای مطالعه‌ی یک قشر اجتماعی مانند پزشکان نمی تواند به همه‌ی آنها مراجعه کند، یا دانش آموزان و دانشجویان یک کشور یا یک شهرستان را نمی تواند به طور کامل مطالعه کند. طبیعی است که انجام دادن تحقیقات در سطح وسیع، اقدام به شکل تمام شماری، نیازمند زمان و امکانات بسیار زیاد می باشد که عملاً مقدور نیست و کار تحقیق را غیر عملی می کند. تنها در سرشماری های نفوس و مسکن آن هم با یک بسیج عمومی از ناحیه‌ی دولت ها است که می توان به افراد جامعه مراجعه کرد.

بنابراین، محقق راه نمونه گیری را انتخاب می کند. محقق به دو شکل ممکن است نمونه را انتخاب کند. یک شکل آن ، این است که شانس انتخاب شدن را به تمامی افراد جامعه بدهد. یعنی روشی را برای انتخاب نمونه برگزیند که تمام افراد جامعه شانس مساوی برای انتخاب شدن داشته باشند. محقق برای انتخاب چنین نمونه ای با در نظر داشتن ویژگی های جامعه‌ی آماری می تواند از روش های گوناگون استفاده کند که به آن اشاره خواهد شد. این روش انتخاب نمونه را روش انتخاب **احتمالی** یا **اتفاقی** می گویند که از آن به روش تصادفی نیز تعبیر می شود. رعایت اصل احتمال و شانس مساوی برای هر یک از افراد جامعه باعث می شود که نمونه‌ی منتخب معرف جامعه باشد و از ارزش علمی برخوردار بوده و صفات آن با صفات جامعه همخوانی و یکنواختی داشته باشد. به این ترتیب، می توان شاخص های محاسبه شده را بر اساس روشهای خاصی به کل جامعه تعمیم داد و پارامترهای آن را برآورد کرد.

روش دیگر برای انتخاب نمونه از بین افراد جامعه، **روش وضعی** و **غیر احتمالی** است. یعنی همه‌ی افراد جامعه شانس مساوی برای انتخاب شدن به عنوان عضو نمونه را ندارند و در انتخاب افراد در نمونه، محقق نظریات خود را دخالت می دهد و به افراد خاصی در جامعه چشم می دوزد. نتایج به دست آمده از چنین مطالعه ای قابلیت تعمیم ندارد. یعنی محقق نمی تواند مقادیر محاسبه شده برای صفات یا شاخص های نمونه

را به جامعه‌ی آماری تعمیم دهد و پارامترهای جامعه را از آن استنباط نماید. اینگونه نمونه‌ها را نمونه‌های **تورش دار** یا **اریب**<sup>۱</sup> نیز می‌گویند.

بنابراین محقق وقتی قصد شناخت جامعه‌ی آماری را از طریق مطالعه‌ی نمونه‌ای از آن دارد باید حتماً روش اول را برای گزینش نمونه انتخاب کند تا نمونه از این طریق با جامعه مشابهت داشته باشد و صفات آنها را بر یکدیگر مطابقت نماید و نمونه معرف جامعه باشد. با توجه به آنچه که اشاره شد، می‌توان تعاریف زیر را بیان کرد.

\*\*\*

**نمونه گیری احتمالی:** نوعی نمونه گیری است که همه‌ی واحدهای آماری شانسی معلوم و مساوی برای انتخاب در نمونه داشته باشند. نمونه گیری احتمالی به روشهای مختلفی انجام می‌گیرد. مهمترین این روش‌ها به شرح زیر هستند.

### الف: نمونه گیری تصادفی ساده

نوعی روش نمونه گیری، که در آن همه‌ی واحدهای آماری برای انتخاب شدن در نمونه و نیز برای گزینش هر یک از نمونه‌های ممکن، شانس یکسان دارند. این روش به شکل‌های مختلف انجام می‌شود. دو روش آن به صورت زیر است:

#### ۱: استفاده از قرئه کشی

در این روش محقق به هر یک از افراد جامعه یک کد یا شماره‌ی مخصوص می‌دهد. سپس از مهره‌ها یا پلاک‌های شماره دار استفاده می‌کند در صورت نبود آن، شماره‌ی هر یک از آنها روی کاغذی و یا مقوای کوچکی یادداشت می‌نماید. بنابراین، به تعداد افراد جامعه، مهره یا پلاک یا کاغذ شماره دار در اختیار خواهد داشت. آنها را در داخل کیسه یا ظرفی می‌ریزد و به هم می‌زند. سپس مهره‌ها را یکی یکی خارج کرده، شماره‌ی آنها را یادداشت می‌نماید و این کار را آنقدر ادامه می‌دهد تا به تعداد حجم نمونه برگزیند. آنگاه که تعداد افراد نمونه کامل شد، کار قرئه کشی به پایان رسیده، مطابق لیست، افراد خود را شناسایی می‌کند.

---

<sup>۱</sup> نمونه‌ای که انتخاب آن احتمالی نباشد یا اندازه‌ی کافی نداشته باشد، نمی‌تواند برآورد دقیقی از جامعه بدهند، چنین نمونه‌هایی را اریب یا تورش دار می‌نامند. اگر نمونه‌ای اریب نباشد، آن را ناریب می‌گویند.

## ۲: استفاده از جدول اعداد تصادفی

در این حالت از اعداد تصادفی برای انتخاب اعضای نمونه استفاده می شود. اعداد تصادفی را می توان از جداول اعداد تصادفی که در کتاب های مختلف آماری، معمولاً درج شده اند، یا از ماشین حساب و یا رایانه آنها را تولید کرد.

**مثال:** یک نمونه تصادفی ۴ تایی از یک جامعه ۱۰ نفره انتخاب کنید.

ابتدا اعضای جامعه را کد گذاری می کنیم.

نام	علی	حسن	مرتضی	زهرا	محمد	فاطمه	مریم	مهدی	رضا	زینب
کد	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰

سپس نمونه گیری را انجام می دهیم.

مرحله	فشار دادن کلید های RAN و SHIFT	ضرب در عدد ۱۰	شماره ی عضو در نمونه (رُند کردن)
اوّل	۰/۳۵۲	۳/۵۲	۴
دوم	۰/۴۵۲	۴/۵۲	۵
سوم	۰/۸۲۵	۸/۲۵	۹
چهارم	۰/۴۸۸	۴/۸۸	-
پنجم	۰/۶۹۱	۶/۹۱	۷

به این ترتیب ۴ عضو نمونه مشخص شدند. این اعضا عبارتند از:

۱: زهرا      ۲: محمد      ۳: رضا      ۴: مریم

## ب: نمونه گیری طبقه ای

نوعی روش نمونه گیری که در آن با طبقه بندی، جامعه به زیر جامعه های مجزاً و همگن تقسیم می شود، سپس از هر طبقه یک نمونه تصادفی انتخاب می شود.

مثال: دانش آموزان پایه ی دهم در رشته تجربی یک روستا به دو طبقه دانش آموزان پسر و دانش آموزان دختر تقسیم می شوند. تعداد آنها طبق جدول زیر است. یک نمونه ۱۰ نفره از این جامعه تعیین کنید.

تعداد کل	دانش آموزان پسر	دانش آموزان دختر
۱۰۰	۴۰	۶۰

ابتدا نسبت افراد نمونه از هر طبقه را تعیین می کنیم.

$$\frac{10}{100} = 0.1$$

تعداد کل	دانش آموزان پسر	دانش آموزان دختر
۱۰	$40 \times 0.1 = 4$	$60 \times 0.1 = 6$

حال از هر طبقه بطور مجزا به یکی از روش های تصادفی ساده یا منظم نمونه انتخاب می کنیم.

**توجه :**

۱: نسبت افراد جامعه در هر طبقه آن باید در نمونه گیری مورد توجه قرار گیرد.

۲: بدیهی است که اگر تعداد افراد طبقات جامعه مساوی باشند، نسبت افراد در نمونه برابر خواهد شد.

### ج: نمونه گیری سامانمند یا سیستماتیک

نوعی روش نمونه گیری طبقه ای است که در آن اندازه‌ی طبقات با هم برابر است و فقط از طبقه‌ی اول واحد های آماری به تصادف انتخاب می شود و با همان رویه از طبقات دیگر، این کار انجام می گیرد.

در این روش همانند روش های قبل فرض بر این است که افراد جامعه متجانس هستند و از این رو به هر یک از آنها از عدد ۱ تا  $N$  شماره (کُد) داده می شود. سپس افراد نمونه با نظمی خاص انتخاب می شوند. این روش نیز ساده است و محققان غالباً از آن استفاده می کنند. در روش منظم محقق سعی می کند، فاصله‌ی عددی دو نمونه را به طور ثابت مشخص کند. آنگاه برای تعیین کد یا شماره‌ی اولین فرد نمونه و مشخص کردن موقعیت آن در سلسله‌ی اعداد و نیز موقعیت سایر افراد نمونه می توان با افزودن یا کاستن عدد ثابت فاصله، اقدام کند. توجه داشته باشید که برای تعیین موقعیت اولین نمونه از روش انتخاب اتفاقی یا احتمالی ساده استفاده می شود. مثلاً بین اعداد ۱ تا ۹ را قرعه کشی کرد و یک عدد را انتخاب نمود. این عدد معرف اولین عضو نمونه خواهد بود. با اضافه کردن فاصله‌ی مورد نظر به کد اولین عضو کد بعدی به دست می آید. این عمل تا انتخاب آخرین عضو ادامه پیدا می کند. به مثال زیر توجه کنید.

**مثال:** محقق می خواهد از بین افراد یک جامعه‌ی دانشجویی ۵۰۰ نفری نمونه ای به تعداد ۵۰ نفر را به روش منظم یا سیستماتیک انتخاب کند. برای این کار پس از کُدگذاری، ابتدا عدد ثابت  $k$  را محاسبه می کند.

$$k = \frac{N}{n} = \frac{500}{50} = 10$$

سپس به روش قرعه کشی موقعیت اولین فرد نمونه را بین اعداد ۱ تا ۹ را مشخص می نماید. فرض کنید که این عدد ۶ انتخاب شود. عدد ۶ کد اولین عضو نمونه است. کد اعضای دیگر به ترتیب زیر است.

۴۹۶ و .... و ۵۶ و ۴۶ و ۳۶ و ۲۶ و ۱۶ و ۶

روش نمونه گیری منظم، باعث می شود تا افراد نمونه به طور یکنواخت در سراسر جامعه پراکنده شوند.

**مثال:** مدیر شهرداری اعلام می کند که در روز های گذشته حدود ۵۰۰ نفر از این شهر بازی دیدن کردند.

برای تعیین میزان رضایت مشتریان، قرار است به میزان ۵۰ فرم پرسش نامه توزیع شود. لذا با توجه به مثال

فوق پرسش نامه به افراد دارای کدهای تعیین شده ( با توجه به قبض ورودی ) انتخاب می شوند.

### د: نمونه گیری خوشه ای

نوعی نمونه گیری که در آن، واحدهای نمونه گیری اولیه در جامعه، گروه ها یا خوشه ها باشند که تفاوت داده ها درون خوشه ها بیشتر از بین خوشه ها است.

**مثال:** محقق می خواهد درباره ی معلمان یا دانش آموزان تحقیقی انجام دهد. با توجه به الگوی تقسیمات

کشوری او باید در مرحله ی اول در سطح استان ها نمونه گیری کند و استان های مورد نظر را انتخاب نماید.

در مرحله ی دوم در بین استان های برگزیده شده، نمونه گیری را در سطح شهرستان انجام دهد. در مرحله ی

سوم باید بین شهرهای انتخاب شده، تعدادی منطقه ی شهری را انتخاب نماید.

از آنجا که واحد تحلیل در تحقیق معلمان است. محقق باید چهارمین مرحله ی نمونه گیری را انجام دهد و در

سطح شهر اقدام به انتخاب چند مدرسه نماید و در نهایت واحد تحلیل و مطالعه ی خود را برگزیند.

همانطور که گفته شد، در هر یک از مراحل نمونه گیری در هر یک از طبقات باید افراد نمونه را از روش

احتمالی ساده یا منظم برگزیده شوند.

**تمرین ۱:** جدول زیر را در مورد روشهای نمونه گیری احتمالی کامل کنید.

معایب	مزایا	روش نمونه گیری
هزینه بر بودن و عدم دسترسی مناسب به واحدهای آماری در جامعه های بزرگ	همه ی اعضا شانس مساوی برای انتخاب دارند.	تصادفی ساده
شانس انتخاب واحد ها متفاوت است.	کم هزینه، سریع	خوشه ای
زمان بر و هزینه بر همچنین شانس انتخاب واحد ها متفاوت است.	اطمینان از انتخاب عضو از همه ی طبقات	طبقه ای
بدون داشتن فهرستی از واحدهای آماری امکان انجام وجود ندارد.	شانس انتخاب یکسان اعضا و بهینه در صرف زمان و هزینه	سیستماتیک



**نمونه گیری غیر احتمالی:** نوعی نمونه گیری است که در آن واحدهای آماری شانسی نامساوی و یا اینکه هیچ شانسی برای انتخاب شدن در نمونه را نداشته باشند. برخی از روش های نمونه گیری های غیر احتمالی عبارتند از:

**۱: نمونه گیری در دسترس (اتفاقی):** فقط از واحدهای آماری که در دسترس هستند، استفاده می کنیم.

**مثال:** این روش یکی از ساده ترین روش ها است. یعنی اینکه افراد مورد مطالعه قرار می گیرند که در دسترس قرار دارند و مصاحبه گر در چهارچوب تعداد و حجم نمونه در مکانهای خاصی می ایستد و با هرکس از راه رسید مصاحبه می کند. مانند خبرنگاران رسانه ها، این مکان ها ممکن است، ایستگاه اتوبوس، بازار، فروشگاه بزرگ، سلف سرویس، رستوران، زیارتگاه، ورودی شهر و .... باشد. این نمونه ها را نمی توان تعمیم داد.

مثال: در فرودگاه مستقر می شویم و در مورد سفرهای نوروزی نظرخواهی می کنیم.

**۲: نمونه گیری قضاوتی:** فقط از گروه خاصی از جامعه، نمونه بگیریم.

مثال:

گاهی اوقات محقق بر اساس تجربه شخصی یا تجارب تکراری و مشابه دیگران یک گروه اجتماعی را معرفی جامعه ای که به آن تعلق دارند می یابد. مثلاً در بعد افکار و عقاید، آنها را منعکس کننده افکار جامعه می یابد و مورد مطالعه قرار می دهد. یا اینکه آموزگاری با شناخت خود از دانش آموزان اقدام به گروه بندی آنها برای آزمایش یک روش تدریس می نماید.

مثال: فقط از قشر فرهنگی پیرامون رضایت شغلی مصاحبه می کنیم.

**۳: نمونه گیری سهمیه ای:** برخی از اعضاء جامعه و یا بخشهای آن، سهم مشخصی در نمونه داشته باشند.

**مثال:** در این روش تعداد نمونه ها مشخص می شود و به همراه دستور العمل مصاحبه تحویل پرسشگر می گردد تا شخصاً به میدان بررسی رفته و خودش افراد نمونه را با توجه به تعدادی که به وی داده شده انتخاب

کند و از طریق مصاحبه‌ی با آنها اطلاعات لازم را گردآوری نماید. این روش هر چند مورد حمایت عده‌ای قرار گرفته، چون در آن اصل شانس برابر برای کلیه‌ی افراد جامعه رعایت نمی‌شود، ارزش علمی مطلوب ندارد و نمی‌توان به تعمیم نتایج آن اعتماد کرد. مانند نمونه‌گیری برای پیش‌بینی اینکه کدام نامزد انتخابات رأی می‌آورد.

**مثال:** برای بررسی میزان رأی نامزدهای انتخابات قبل از رأی‌گیری، در نزدیکی ستاد یکی از نامزدهای انتخاب مستقر و به مدت ۲ ساعت پرسش می‌کنیم ولی در نزدیکی ستاد دیگر فقط ۱۵ دقیقه پرسش می‌کنیم.

\*\*\*

**تمرین ۲:** فرض کنید جامعه‌ای از ۱۰۰ عضو تشکیل شده است و می‌خواهیم نمونه‌ای به اندازه‌ی ۲۰ نفر از آن انتخاب کنیم. در هر یک از حالت‌های زیر، احتمال انتخاب هر عضو جامعه به عنوان نمونه چقدر است؟ نام هر روش نمونه‌گیری را بنویسید.

الف: اگر جامعه به دو قسمت ۵۰ تایی تقسیم شود و بخواهیم از هر قسمت، نمونه‌ی تصادفی ۱۰ تایی انتخاب کنیم.

ب: اگر جامعه به تصادف به ۱۰ قسمت مساوی تقسیم شود و دو قسمت را به عنوان نمونه انتخاب کنیم.

پ: اگر جامعه به تصادف به ۲۰ قسمت مساوی تقسیم شود و از قسمت اول یک عضو به تصادف انتخاب شود. فرض کنید عضو انتخابی دومین عضو باشد و از قسمت‌های بعدی نیز دومین عضو انتخاب شود.

حل:

احتمال	نام روش	
$\frac{10}{50} = \frac{1}{5}$	طبقه‌ای	الف
$\frac{10}{100} = \frac{1}{10}$	خوشه‌ای	ب
احتمال اولی $\frac{1}{5}$ و بقیه ۱ است.	سیستماتیک	پ

**تمرین ۳:** در هر مثال دلیل آریبی را ذکر کنید.

الف: نمونه گیری از اعضای تیم ملی والیبال برای بررسی درآمد مالی والیبالیست ها در ایران

ب: نمونه گیری از اهالی نزدیک ترین محله در تهران برای بررسی سطح رفاه اجتماعی در این شهر

حل:

الف: انتخاب بخش و گروه خاصی از جامعه‌ی والیبال ایران و عدم وجود شانس انتخاب برای واحدهای آماری

ب: فقط از افراد در دسترس استفاده شده است و همه‌ی اعضا شانس انتخاب ندارند.

**تمرین ۴:** آیا در نمونه گیری خوشه ای، احتمال انتخاب واحدهای آماری برابر است؟ چرا؟

جواب: در روش خوشه ای، ابتدا جامعه به گروه ها یا خوشه هایی تقسیم می شوند. سپس از بین خوشه ها

تعدادی به تصادف انتخاب شده و سپس همه‌ی واحدهای آن مورد مطالعه قرار می گیرند. پس خوشه ها از

شانس برابری برای انتخاب برخوردار هستند، اما واحدهای آماری از شانس برابر برخوردار نیستند. البته اگر تعداد

اعضای خوشه ها با هم برابر باشند، واحدهای آماری می توانند شانس برابر داشته باشند، اما در حالت کلی این

مورد برقرار نیست.

**تمرین برای حل:**

**۵:** جای خالی را کامل کنید.

الف: هر زیر مجموعه از جامعه‌ی آماری که به روش مشخصی انتخاب شده باشد را یک ..... می نامند.

ب: نوعی روش نمونه گیری که در آن، واحدهای آماری برای انتخاب شدن در نمونه، احتمال یکسان دارند،

نمونه گیری ..... است.

پ: ۱۰۰ دانش آموز را به ۵ گروه مساوی تقسیم کرده و از میان این ۵ گروه، ۲ گروه را که جمعاً ۴۰ نفر

هستند به اردو می برند. در انتخاب این ۴۰ نفر از نمونه گیری ..... استفاده شده است.

ت: نمونه هایی نمی تواند برآورد دقیقی از جامعه بدهند، را ..... می نامند.

\*\*\*

## روش‌های گردآوری داده‌ها

داده‌ها واقعیت‌هایی درباره‌ی یک موضوع هستند که در محاسبه، استنباط یا برنامه‌ریزی به کار می‌روند. عمل گردآوری داده‌ها به یکی از روش‌های ممکن را **آمارگیری** و کسی که آمارگیری انجام می‌دهد را **آمارگیر** می‌نامند. برای جمع‌آوری داده‌ها روش‌های مختلفی وجود دارد. معمول‌ترین این روش‌ها عبارتند از:

**(۱) مشاهده:** گردآوری داده‌ها بدون نیاز به فرد پاسخ‌گو

مثال: شمارش تعداد وسایل نقلیه‌ی عبوری از یک تقاطع در هر ساعت

**(۲) پرسش از واحد‌های آماری**

**کتبی (پرسش‌نامه):** مجموعه‌ی سؤال‌های از پیش تعیین شده که توسط تعدادی پاسخ‌دهنده تکمیل می‌شود.

مثال: پرسش‌نامه‌ی مربوط به رضایت معلمان پیرامون یک دوره‌ی ضمن خدمت که برگزار شده است.

**شفاهی (مصاحبه):** حالت خاصی از پرسش‌نامه که به صورت شفاهی صورت می‌گیرد.

مثال: مصاحبه پیرامون راه حل مناسب برای برطرف نمودن مشکل ترافیک شهر

**(۳) دادگان:** استفاده از داده‌های از پیش تهیه شده که در گذشته جمع‌آوری شده‌اند و می‌توان از آن‌ها در تحقیقات بعدی بهره برد.

مثال: بررسی روند تغییر نمرات ریاضی یک دبیرستان از ده سال گذشته تاکنون

**(۴) آزمایش:** جمع‌آوری اطلاعات از واحد‌های آماری با انجام یکسری آزمایش روی آن‌ها

این روش برای علوم تجربی بسیار مرسوم است.

مثال: بررسی اثر خورشید بر پوست انسان

**تمرین ۶:** بهترین روش جمع‌آوری داده‌ها را برای موارد زیر بیان کنید.

(۱) میزان رضایت مشتریان بانک از نحوه‌ی برخورد و رسیدگی به درخواست‌های آن‌ها

(۲) سن همه‌ی دانش‌آموزان بر حسب ماه در پایه‌ی دهم

(۳) تعداد سرنشینان خودروهای سواری در یکی از محورهای خروجی شهر

۴) بررسی رشد جمعیت در ۵۰ سال گذشته‌ی جهان

۵) بررسی سطح کیفی خدمات بانکی خاص در یک محله

۶) بررسی شرایط متقاضیان وام در یک بانک

۷) بررسی تعداد تصادف در خیابانی در مشهد

۸) بررسی میزان چربی خون افراد در سنین نوجوانی

حل :

۱) پرسش نامه      ۲) دادگان      ۳) مشاهده      ۴) دادگان      ۵) مصاحبه      ۶) پرسش نامه

۷) مشاهده      ۸) آزمایش

\*\*\*

## آمار استنباطی

هر ویژگی از اشخاص یا اشیاء که قرار است بررسی شود را **متغیر** می نامند. ویژگی اصلی متغیر این است که

از هر عضو جامعه یا نمونه به عضو دیگر تغییر می کند. متغیرها دارای دو نوع اصلی می باشد، کمی و کیفی

**متغیر کمی** : متغیرهایی هستند که مقادیر عددی می گیرند و برای آنها عملیات ریاضی از قبیل جمع ،

تفریق و معدل گیری قابل انجام است.

مثال : سن ، نمره ، تعداد افراد خانواده ، طول مکالمه ی تلفنی و ...

**متغیر کیفی** : متغیرهایی هستند که صرفاً برای دسته بندی افراد یا اشیاء در گروه ها به کار می روند و لزوماً

مقدار عددی نمی گیرند.

مثال : جنسیت ( مرد و زن ) ، کیفیت کالا ( مرغوب ، متوسط ، نامرغوب )

## تمرین برای حل :

۷: جای خالی را کامل کنید.

الف : به هر یک از ویژگی هایی که مورد پژوهش قرار می گیرند، ..... می گویند.

ب : میزان بارندگی یک متغیر ..... است.

ج : گروه خونی یک متغیر ..... است.

## شاخص های نمونه و جامعه

بعد از جمع آوری اطلاعات از جامعه یا نمونه می توان، شاخص هایی در مورد آنها را محاسبه کرد. بستگی به این که این شاخص ها از جامعه یا از نمونه به دست آمده باشند به دو دسته ی پارامتر و آماره، تقسیم می شوند.

هر مشخصه ی عددی که توصیف کننده ی جنبه ی خاصی از جامعه است را پارامتر می نامند. در محاسبه ی پارامتر، داده های تمامی اعضای جامعه در نظر گرفته می شوند.  
مانند:

الف: نسبت کل معلمان ورزشکار شهرستان به کل معلمان این شهرستان

ب: میانگین سن کل معلمان ورزشکار شهرستان

هر مشخصه ی عددی که توصیف کننده ی جنبه ای خاص از نمونه باشد، را آماره می نامند. در محاسبه ی آماره فقط داده های اعضای نمونه شرکت داده می شوند.  
مانند:

الف: نسبت معلمان ورزشکار در بین معلمان دو دبیرستان شهرستان که به تصادف انتخاب می شوند.

ب: میانگین سن معلمان ورزشکار در بین معلمان دو دبیرستان شهرستان که به تصادف انتخاب می شوند.

پارامترها و آماره های مهم که در آمار مورد استفاده قرار می گیرند، به شرح جدول زیر هستند.

شاخص	گروه	نماد کلی	نسبت	میانگین	واریانس
آماره	نمونه	$\hat{\theta}$	$p$	$\bar{x}$	$S^2$
پارامتر	جامعه	$\theta$	$P$	$\mu$	$\sigma^2$

توجه: چون در محاسبه ی پارامترها تمامی داده های جامعه در نظر گرفته شده اند لذا پارامترها در جامعه ثابت هستند. ولی آماره ها در نمونه های مختلف از یک جامعه، متفاوت می باشند. بدیهی است که هر چه حجم نمونه به حجم جامعه نزدیکتر شود، آماره به پارامتر نزدیک می شود.

فرآیند نتیجه گیری درباره ی پارامترهای جامعه بر اساس نمونه، آمار استنباطی نامیده می شود.

**مثال ۱:** در یک مطالعه، ۱۲۶۱ مشتری غذاهای گیاهی یک شهرستان، سؤال شده است که برای کدام

وعده ی غذایی (ناهار یا شام) سفارش داده اند. با توجه به این موضوع به سئوالات زیر پاسخ دهید.

الف: متغیر را مشخص کنید و نوع آن را بنویسید.

**جواب:** متغیر، وعده‌ی غذایی و از نوع کیفی می باشد.

ب: کدام روش گردآوری داده ها برای این مطالعه مناسب است؟

**جواب:** پرسش نامه

ج: جامعه‌ی آماری در این جا چیست؟ در این مطالعه پارامتر و آماره چه چیزی می توانند باشند؟

**جواب:** جامعه‌ی آماری مشتریان غذای گیاهی شهرستان مورد نظر است. پارامتر و آماره نسبت سفارشات نهار

به کل سفارشات می باشند.

**مثال ۲:** در بررسی میانگین درآمد خانوارهای کشور، از هر استان، ۲۰ خانواده انتخاب کرده و میانگین آنها را

به دست می آوریم. با توجه به این مطالعه سئوالات زیر را پاسخ دهید.

الف: متغیر مورد مطالعه و نوع آن را بنویسید؟

**جواب:** درآمد خانوارهای کشور، کمی

ب: در این مطالعه، نمونه و حجم آن را تعیین کنید.

**جواب:** خانوارهای تعیین شده، حجم آن ۲۰ خانوار است.

ج: پارامتر و آماره را بنویسید.

**جواب:** میانگین درآمد خانوارهای کشور، پارامتر و میانگین درآمد خانوارهای خانواده های انتخاب شده،

آماره است.

**مثال ۳:** برای تعیین میانگین درآمد خانوارهای یک روستا در استان کرمان، ۶ خانوار به تصادف با درآمدهای

۳/۵ و ۱/۵ و ۳ و ۵ و ۴ و ۱ میلیون تومان انتخاب شده اند. آماره‌ی میانگین نمونه را به دست آورید.

$$\bar{x} = \frac{1 + 4 + 5 + 3 + 1/5 + 3/5}{6} = 3$$

**تمرین برای حل:**

**۸:** جای خالی را کامل کنید.

الف: در یک جامعه‌ی آماری ..... ثابت است ولی آماره بین نمونه های مختلف جامعه ..... می کند.

ب : هرچه نمونه، تصادفی انتخاب شده و به اندازه‌ی کافی بزرگ باشد، در این صورت ..... به پارامتر نزدیک می شود.

پ : فرآیند نتیجه گیری درباره‌ی پارامترهای جامعه بر اساس نمونه، ..... نامیده می شود.

ت : تفاوت در نوع و کاربردهای آماره و پارامتر ، موجب تقسیم بندی آمار به آمار توصیفی و آمار ..... شده است. آمار توصیفی به خلاصه کردن داده ها می پردازد و هدف آن بررسی داده های گردآوری شده است ولی آمار استنباطی به تجزیه و تحلیل نمونه و ..... آن به جامعه می پردازد.

\*\*\*

---



## درس ۲: برآورد (تخمین)

علم آمار بر اساس میزان اطلاعات در دسترس از نمونه یا جامعه به دو بخش آمار توصیفی و آمار استنباطی تقسیم می شود. آمار توصیفی به خلاصه کردن داده های نمونه می پردازد. به کمک اطلاعات این نمونه ها، می توانیم به برآوردی درست از ویژگی های جامعه برسیم. در این حالت از آمار استنباطی به تجزیه و تحلیل نمونه و بررسی ویژگی های جامعه می پردازد.

فرض کنید می خواهیم میزان رضایت افراد جامعه از زندگیشان را بررسی کنیم. برای بررسی لازم هست از پارامترهای مرکزی و پراکندگی در جامعه استفاده کنیم. اگر اطلاعات همه ی افراد جامعه را داشته باشیم، می توان از پارامترها استفاده کرد. اما در اغلب موارد اطلاعات کامل جامعه را نداریم. پس مجبوریم از نمونه گیری تصادفی استفاده کنیم. پس نمونه گیری به کمک آماره ی مورد نظر « آماره جایگزین نمونه است.» می توان تخمینی از مقادیر پارامتر مربوط به آن داشت. بر این اساس می توان به طور خلاصه نوشت:

آمار توصیفی : خلاصه کردن داده ها

آمار استنباطی : تجزیه و تحلیل نمونه و تعمیم آن به جامعه

برای تحلیل یک پارامتر مجهول در جامعه، باید نمونه ای از جامعه را انتخاب کرده و بر اساس نتایج به دست آمده از روی نمونه با انتخاب یک آماره ی مناسب ( جایگزین پارامتر در نمونه) به بحث و نتیجه گیری در مورد پارامتر مورد نظر پرداخت. یکی از روشهای انجام این عمل « برآورد» پارامتر مجهول جامعه است. برآورد به دوشکل انجام می گیرد.

## الف : برآورد نقطه ای

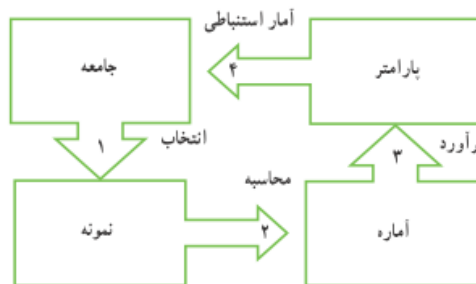
برآورد نقطه ای پارامتر جامعه برابر با مقدار عددی حاصل از جایگذاری اعداد نمونه ی تصادفی در آماره ی نظیر پارامتر است.

$$\text{مثلاً: } \mu_x = \bar{x} = 15/2$$

واضح است که احتمال صحت برآورد نقطه ای بسیار ناچیز است و شاید بتوان گفت برابر صفر است. اما چون برآورد نقطه ای از نمونه ی تصادفی تبعیت می کند، برای هر پارامتر مجهول ، برآوردهای مختلف به ازای

نمونه های مختلف به دست می آید. اما توجه به این نکته مهم است که **با زیاد شدن اندازهی نمونه،**

**فاصله‌ی آماره تا پارامتر کمتر می شود.** لذا برآورد دقیق تری خواهیم داشت.



به انحراف معیار میانگین های نمونه های جامعه، خطای معیار میانگین گفته می شود. خطای معیار میانگین، شاخص پراکندگی میانگین های نمونه ها از میانگین جامعه است.

هر چقدر مقدار خطای معیار میانگین، کمتر باشد، میانگین نمونه ها با خطای کمتری، میانگین جامعه را نشان می دهد و برآورد دقیق تری از میانگین جامعه خواهد داشت.

خطای معیار میانگین، میزان برآورد از مقدار واقعی میانگین را نشان می دهد. به خطای معیار میانگین، انحراف معیار برآورد میانگین هم گفته می شود.

آمار دانان انحراف معیار میانگین را به صورت زیر محاسبه می کنند.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

که در آن  $\sigma$  انحراف معیار جامعه،  $n$  حجم نمونه و  $\sigma_{\bar{x}}$  انحراف معیار میانگین می باشد.

با توجه به این رابطه واضح است که هر چقدر  $n$  افزایش یابد، با توجه به ثابت بودن  $\sigma$  در جامعه، مقدار خطای میانگین ها کاهش می یابد.

**مثال:** نتیجه‌ی یک آزمون در سطح دبیرستان های یک استان به شرح زیر است.

میانگین ۵۰ و انحراف معیار ۱۴، فرض کنید از جامعه، چندین گروه تصادفی با حجم  $n = ۴۹$  انتخاب و

میانگین هر یک از گروه ها محاسبه شده است. خطای معیار میانگین را محاسبه کنید.

حل:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{۱۴}{\sqrt{۴۹}} = \frac{۱۴}{۷} = ۲$$

## ب: برآورد فاصله ای

گفتیم برای بررسی پارامترهای جامعه، از آن نمونه‌ی تصادفی گرفته و به کمک آماره‌ی مناسب، پارامتر را برآورد نقطه ای می کنیم. چون برآورد بر اساس داده های موجود در نمونه صورت گرفته است، امکان خطا در آن وجود دارد. این خطا، به میزان تفاوت در میانگین نمونه های انتخاب شده بستگی دارد که با اصل افزایش تعداد اعضای نمونه، مقدار خطا کاهش می یابد.

میانگین نمونه (آماره)، بهترین برآورد نقطه ای میانگین جامعه می باشد ولی امکان خطا در آن زیاد است. باتوجه به این توضیحات لزوم ارائه‌ی روشی دیگر، غیر از روش برآورد نقطه ای وجود دارد. به این برآورد، برآورد بازه ای یا فاصله ای گفته می شود.

اگر دقت برآورد یا خطای نمونه گیری را با  $\epsilon$  نمایش دهیم، برآورد فاصله ای به صورت زیر بیان می شود.

$$\epsilon \pm \text{آماره}$$

در برآورد فاصله ای، فاصله‌ی بدست آمده را فاصله‌ی اطمینان و حدود بدست آمده را حدود اطمینان می نامند. در ادامه چند مورد از برآورد های جامعه را معرفی می کنیم.

## برآورد فاصله ای میانگین جامعه

برآورد بازه ای میانگین جامعه عبارت است از بازه های عددی برای پارامتر میانگین به همراه یک درصد اطمینان که به ضریب اطمینان شهرت دارد. اگر نمونه ای تصادفی به اندازه‌ی  $n$  در اختیار داشته باشیم، با اطمینان بیش از ۰/۹۵ می توانیم بگوییم:

$$\bar{x} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{2\sigma}{\sqrt{n}}$$

که در آن  $\mu$  میانگین جامعه،  $\sigma$  انحراف معیار جامعه،  $n$  حجم نمونه و  $\bar{x}$  میانگین نمونه است.

**مثال:** فرض کنید  $\bar{x} = 50$ ،  $\sigma = 20$  و  $n = 100$ . بازه‌ی اطمینان میانگین جامعه را محاسبه کنید.

حل:

$$\bar{x} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{2\sigma}{\sqrt{n}} \rightarrow 50 - \frac{2 \times 20}{\sqrt{100}} < \mu < 50 + \frac{2 \times 20}{\sqrt{100}} \rightarrow 46 < \mu < 54$$

**مثال :** یک نمونه‌ی تصادفی از ۶۴ لامپ، نشان می‌دهد که عمر متوسط نمونه ۳۵۰ ساعت است. فاصله‌ی اطمینان ۹۵ درصد، برای متوسط طول عمر واقعی لامپ‌ها با فرض  $\sigma_x = 100$  را به دست آورید.

حل :

$$\bar{x} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{2\sigma}{\sqrt{n}} \rightarrow 350 - \frac{2 \times 100}{\sqrt{64}} < \mu < 350 + \frac{2 \times 100}{\sqrt{64}} \rightarrow 325 < \mu < 375$$

**مثال :** در یک جامعه، می‌دانیم میانگین برآورد شده توسط نمونه، ۴ واحد با میانگین جامعه اختلاف دارد.

الف : اگر نمونه ۴۰۰ تایی باشد، حداقل انحراف معیار جامعه را به دست آورید.

حل :

$$\bar{x} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{2\sigma}{\sqrt{n}} \rightarrow |\mu - \bar{x}| < \frac{2\sigma}{\sqrt{n}} \rightarrow 4 < \frac{2\sigma}{\sqrt{400}} \rightarrow 2 < \frac{\sigma}{20} \rightarrow \sigma > 40$$

ب : اگر واریانس جامعه ۱۰۰ واحد باشد، حداکثر حجم نمونه را بیابید.

حل :

$$\bar{x} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{2\sigma}{\sqrt{n}} \rightarrow |\mu - \bar{x}| < \frac{2\sigma}{\sqrt{n}} \rightarrow 4 < \frac{2 \times 100}{\sqrt{n}} \rightarrow \sqrt{n} < 50 \rightarrow n \leq 2500$$

**نتیجه :** با توجه به برآورد فاصله‌ی میانگین جامعه، می‌توان گفت که دقت برآورد میانگین جامعه به اندازه‌ی نمونه و انحراف معیار جامعه بستگی دارد. اگر اندازه‌ی نمونه زیاد شود یا انحراف معیار جامعه کم شود، دقت برآورد بیشتر می‌شود.

**مثال :** در برآورد فاصله‌ی میانگین، طول بازه‌ی اطمینان را محاسبه کنید. سپس جای خالی را در هر یک از جملات زیر کامل کنید.

ب : اگر اندازه‌ی نمونه افزایش یابد، طول بازه‌ی اطمینان ..... می‌یابد.

ج : اگر انحراف معیار جامعه افزایش یابد، طول بازه‌ی اطمینان ..... می‌یابد.

$$\bar{x} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{2\sigma}{\sqrt{n}} \quad \text{فاصله‌ی اطمینان}$$

$$\rightarrow \left(\bar{x} + \frac{2\sigma}{\sqrt{n}}\right) - \left(\bar{x} - \frac{2\sigma}{\sqrt{n}}\right) = \bar{x} + \frac{2\sigma}{\sqrt{n}} - \bar{x} + \frac{2\sigma}{\sqrt{n}} = \frac{4\sigma}{\sqrt{n}} \quad \text{طول بازه‌ی اطمینان}$$

الف : کاهش ، زیرا بین طول بازه‌ی اطمینان با جذر اندازه‌ی نمونه رابطه‌ی معکوس وجود دارد.

ب : افزایش ، زیرا بین طول بازه‌ی اطمینان با انحراف معیار جامعه رابطه‌ی مستقیم وجود دارد.

## برآورد فاصله نسبت جامعه

از ساده ترین مسائل در آمار استنباطی، برآورد کردن یک نسبت در جامعه است. مثلاً برآورد نسبتی از مردم که می خواهند در تعطیلات نوروز سال آینده به مسافرت بروند. یا درصدی از کشاورزان که در سال گذشته از آبیاری قطره ای در مزارع خود استفاده کرده اند و ...

با توجه به اهمیت این مسائل در برنامه ریزیها و برای برآورد کردن درست این نسبت ها، از برآورد فاصله ای نسبت استفاده می شود. این برآورد مانند برآورد میانگین جامعه با اطمینان ۹۵ درصد انجام می شود.

اگر از جامعه ای نسبتاً بزرگ،  $n$  نمونه ای تصادفی ساده انتخاب کنیم و  $m$  تا از آنها ویژگی مورد مطالعه ای ما را داشته باشد، آنگاه نسبت واقعی افرادی از جامعه که آن ویژگی را دارند، با اطمینان ۹۵ درصد بازه ای زیر است.

$$p - 2\sqrt{\frac{p(1-p)}{n}} < P < p + 2\sqrt{\frac{p(1-p)}{n}}$$

که در آن  $p = \frac{m}{n}$  برآورد پارامتر نسبت ویژگی در جامعه است.

**مثال:** فرض کنید، از ۴۸ دانش آموز مدرسه پرسیده ایم که، آیا برای آمدن به مدرسه، از وسایل نقلیه ای عمومی استفاده می کنید؟ از بین این دانش آموزان به تعداد ۳۶ نفر به سؤال ما جواب مثبت داده اند. در این صورت، تعیین کنید که چند درصد از دانش آموزان مدرسه جوابشان به این سؤال مثبت خواهد شد؟ (با اطمینان ۹۵ درصد)

حل: از صورت مسئله معلوم می شود که تعداد نمونه ۴۸ است و ۳۶ نفر این ویژگی را دارند. یعنی  $n = 48$

و  $m = 36$ ، پس  $p = \frac{36}{48} = 0.75$  لذا داریم:

$$p - 2\sqrt{\frac{p(1-p)}{n}} < P < p + 2\sqrt{\frac{p(1-p)}{n}}$$

$$\rightarrow 0.75 - 2\sqrt{\frac{0.75(1-0.75)}{48}} < P < 0.75 + 2\sqrt{\frac{0.75(1-0.75)}{48}}$$

$$\rightarrow 0.75 - 0.125 < P < 0.75 + 0.125 \rightarrow 0.625 < P < 0.875 \rightarrow \frac{63}{100} < P < \frac{88}{100}$$

یعنی با اطمینان ۹۵ درصد، نسبت دانش آموزان مدرسه که از وسایل نقلیه ای عمومی استفاده می کنند در

بازه ای (۰/۶۳, ۰/۸۸) قرار دارد.

**توجه:** در برآورد بالا، طول بازه‌ی اطمینان ۰/۹۵ می باشد. حال اگر تعداد نمونه‌ها را  $k$  برابر کنیم. طول بازه‌ی اطمینان بر  $\sqrt{k}$  تقسیم می شود. پس اگر تعداد نمونه‌ها را ۴ برابر کنیم، طول بازه‌ی اطمینان نصف می شود و اگر تعداد نمونه‌ها را صد برابر کنیم، دقت محاسبه‌ی نسبت یک رقم اعضا بهتر خواهد شد.

در رابطه‌ی برآورد نسبت،  $p$  در واقع میانگین نسبت ویژگی در مورد مطالعه در نمونه‌ها و  $\sqrt{\frac{p(1-p)}{n}}$  همان انحراف معیار نسبت‌های نمونه‌ای می باشند.

**مثال:** در یک نظر سنجی پیش از انتخابات، فرض کنید که از پیش بدانیم که آرای یکی از نامزدها نزدیک به ۵۰ درصد است. اگر بخواهیم طول بازه‌ی اطمینان ۹۵ درصدی، کمتر از یک درصد باشد، نمونه‌ی ما باید شامل چند نفر باشد؟

حل:

$$p - 2\sqrt{\frac{p(1-p)}{n}} < P < p + 2\sqrt{\frac{p(1-p)}{n}}$$

$$\rightarrow (p + 2\sqrt{\frac{p(1-p)}{n}}) - (p - 2\sqrt{\frac{p(1-p)}{n}}) < \frac{1}{100}$$

$$\rightarrow 2\sqrt{\frac{p(1-p)}{n}} + 2\sqrt{\frac{p(1-p)}{n}} < \frac{1}{100} \rightarrow 4\sqrt{\frac{p(1-p)}{n}} < \frac{1}{100} \rightarrow \sqrt{\frac{p(1-p)}{n}} < \frac{1}{400}$$

طبق مسئله  $p = 0/5$  می باشد. پس:

$$\sqrt{\frac{0/5(1-0/5)}{n}} < \frac{1}{400} \rightarrow \frac{0/5}{\sqrt{n}} < \frac{1}{400} \rightarrow \frac{1}{2\sqrt{n}} < \frac{1}{400} \rightarrow \sqrt{n} > 200 \rightarrow n > 40000$$

**توجه ۱:** عبارت  $p(1-p)$  یک چند جمله‌ای درجه‌ی ۲ است و بیشترین مقدار خود را در حالت

$p = 0/5$  می گیرد. یعنی

$$p(1-p) = 0/5(1-0/5) = (0/5)^2$$

لذا برآورد فاصله این نسبت در این حالت به صورت زیر خواهد شد.

$$p - 2\sqrt{\frac{p(1-p)}{n}} < P < p + 2\sqrt{\frac{p(1-p)}{n}} \rightarrow p - 2\sqrt{\frac{(.05)^2}{n}} < P < p + 2\sqrt{\frac{(.05)^2}{n}}$$

$$\rightarrow p - 2 \times \frac{1}{2} \sqrt{\frac{1}{n}} < P < p + 2 \times \frac{1}{2} \sqrt{\frac{1}{n}} \rightarrow p - \frac{1}{\sqrt{n}} < P < p + \frac{1}{\sqrt{n}}$$

لذا در برآورد فاصله ای نسبت، می توان گفت، با اطمینان بیشتر از ۹۵ درصد، نسبت مورد نظر در این بازه زیر است.

$$p - \frac{1}{\sqrt{n}} < P < p + \frac{1}{\sqrt{n}}$$

که در آن  $p = \frac{m}{n}$  می باشد.

**توجه ۲:** در برآورد بازه ای برای نسبت، دقت یا خطا عبارت است از

$$e = 2\sqrt{\frac{p(1-p)}{n}}$$

می باشد. این مقدار میزان حداکثر خطای مجاز را نشان می دهد.

**مثال:** مدیر تولید یک روزنامه می خواهد درصد روزنامه های معیوب را بررسی کند. برای این منظور

۱۰۰ روزنامه به تصادف انتخاب می شود که ۱۶ تای آنها معیوب است. یک فاصله ای اطمینان ۰/۹۵ برای

روزنامه های معیوب محاسبه کنید.

حل: طبق مسئله  $p = \frac{16}{100}$  می باشد. حال داریم.

$$p - 2\sqrt{\frac{p(1-p)}{n}} < P < p + 2\sqrt{\frac{p(1-p)}{n}}$$

$$\rightarrow 0.16 - 2\sqrt{\frac{0.16 \times 0.84}{100}} < P < 0.16 + 2\sqrt{\frac{0.16 \times 0.84}{100}} \rightarrow 0.16 < P < 0.23$$

**مثال:** با توجه به مثال قبل، اگر بخواهیم، طول بازه ای اطمینان برابر یک درصد باشد، باید  $n$  را چقدر

انتخاب کنیم.

حل: طبق مثال های قبل برای این حالت داریم.

$$\sqrt{\frac{p(1-p)}{n}} = \frac{1}{400} \xrightarrow{p=0.16} \sqrt{\frac{0.16 \times 0.84}{n}} = \frac{1}{400} \rightarrow \sqrt{\frac{0.1344}{n}} = \frac{1}{400}$$

$$\rightarrow \frac{0.1344}{n} = \frac{1}{160000} \rightarrow n = 21504$$

**مثال:** به منظور ارزیابی کیفیت محصولات تولید شده، تعداد ۲۰۰ واحد محصول را به طور تصادفی انتخاب کرده ایم که بین آنها ۴۰ محصول نقص دار مشاهده شده است. دقت تخمین یا حداکثر خطای مجاز با اطمینان ۹۵ درصد را به دست آورید.

حل: با توجه به مسئله داریم:  $m = 40$  و  $n = 200$  لذا

$$p = \frac{40}{200} = \frac{1}{5}$$

پس حداکثر خطای مجاز به صورت زیر است.

$$e = 2 \sqrt{\frac{p(1-p)}{n}} = 2 \sqrt{\frac{\frac{1}{5} \times \frac{4}{5}}{200}} = \frac{4}{50\sqrt{2}} = \frac{\sqrt{2}}{25}$$

**مثال:** حداقل حجم نمونه‌ی مناسب برای تخمین نسبت افرادی که در انتخابات آینده شرکت می‌کنند، با خطای ۲ درصد و ضریب اطمینان ۹۵ درصد تقریباً چقدر است؟

حل: زمانی که اطلاعات ما در مورد نسبت مورد نظر، محدود است، مقدار  $p$  را  $0.5$  در نظر می‌گیریم و برای تخمین  $n$  بر اساس آن محاسبات را انجام می‌دهیم.

$$e = 0.02 \rightarrow 2 \sqrt{\frac{p(1-p)}{n}} = 0.02 \xrightarrow{p=0.5} 2 \sqrt{\frac{\frac{1}{2} \times \frac{1}{2}}{n}} = \frac{1}{50} \rightarrow \frac{1}{\sqrt{n}} = \frac{1}{50}$$

$$\rightarrow \sqrt{n} = 50 \rightarrow n = 2500$$

لذا حداقل حجم نمونه‌ی مناسب ۲۵۰۰ واحد آماری باید باشد.

\*\*\*

### تمرین برای حل:

۱: جای خالی را کامل کنید.

الف: مقدار عددی حاصل از جایگذاری اعداد نمونه‌ی تصادفی در آماره‌ی نظیر پارامتر را برآورد ..... پارامتر جامعه می‌نامند.

ب: در صورتی که اندازه‌ی نمونه را کمتر کنیم، امکان نزدیک شدن برآورد به پارامتر ..... می‌شود.



پ: اگر تعداد نمونه ها را ۹ برابر کنیم. طول بازه‌ی اطمینان نسبت بر ..... تقسیم می شود.

ت: زمانی که اندازه‌ی نمونه با ..... برابر باشد، آماره با پارمتر برابر می شود.

ث: انحراف معیار برآورد میانگین، از تقسیم ..... بر ..... به دست می آید.

**۲:** در یک اداره که انحراف معیار سن همه‌ی کارمندان آن ۲۰ است، نمونه‌ی ۱۶ تایی به طور تصادفی انتخاب می کنیم که میانگین سن آنها ۴۰ است. یک فاصله‌ی اطمینان ۹۵ درصدی برای میانگین سن همه‌ی کارمندان این اداره به دست آورید.

**۳:** در یک نمونه‌ی ۱۰۰ تایی از جامعه، میانگین برابر ۱۰ و ضریب تغییرات ۲ است. اگر بدانیم انحراف معیار جامعه با انحراف معیار نمونه برابر است، یک فاصله‌ی اطمینان ۹۵ درصدی برای میانگین جامعه به دست آورید و طول این بازه را مشخص کنید.

**۴:** اگر از یک کلاس ۳۶ نفری، ۴ نفر را به عنوان نمونه انتخاب کنیم و نمرات در آمار و احتمال آنها برابر ۱۲ و ۱۶ و ۱۴ و ۱۰ باشد و اگر بدانیم واریانس جامعه برابر ۱۷ باشد. یک فاصله‌ی اطمینان ۹۵ درصدی برای میانگین نمرات درس آمار و احتمال این کلاس بیابید.

**۵:** در نمونه‌گیری از یک جامعه‌ی بزرگ، اگر میانگین ۲۲ و واریانس ۴ برابر اندازه‌ی نمونه باشد. با ضریب اطمینان ۹۵ درصدی، حد بالا و پایین فاصله‌ی اطمینان برای میانگین را به دست آورید. (اندازه‌ی نمونه بیشتر از ۴۰ نفر است).

**۶:** در یکی از محله‌های مشهد از میان ۱۰۰ نفر که به تصادف انتخاب شده‌اند، ۲۰ نفر زائر غیر مشهدی هستند. یک فاصله‌ی اطمینان ۹۵ درصد برای زائرین غیر مشهدی پیدا کنید.

\*\*\*