

به نام خدا

بدین وسیله گواهی می‌شود جناب آقای دکتر مهدی جباری نوقابی در

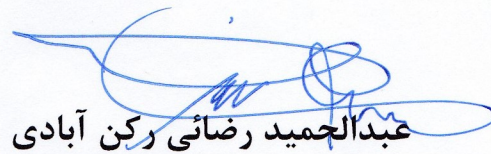
دومین سمینار آموزش آمار

که در تاریخ‌های ۸ و ۹ آبان ماه ۱۳۹۸ در دانشکده علوم ریاضی دانشگاه  
فردوسی مشهد برگزار گردید، شرکت کرده و مقاله خود را تحت عنوان

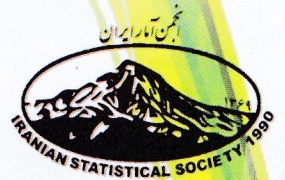
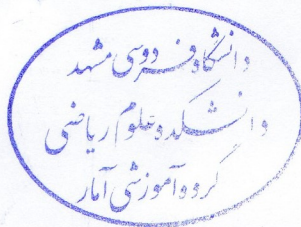
**آموزش آمار و اهمیت آن در تحقیقات پزشکی**

را به صورت سخنرانی ارائه نمودند.

نویسندگان همکار: محیا قاسمی، نسرین تلخی

  
عبدالحامید رضائی رکن آبادی

دبیر سمینار



نشانی دبیرخانه:

استان خراسان رضوی

مشهد - میدان آزادی

پردیس دانشگاه فردوسی مشهد

دانشکده علوم ریاضی

صندوق پستی: ۱۱۵۹ - ۹۱۷۷۵

تلفن: ۰۵۱ - ۳۸۸۰۶۶۰۶

وبگاه: <https://sse2.um.ac.ir>



## آموزش آمار و اهمیت آن در تحقیقات پزشکی

محیا قاسمی<sup>۱</sup>، نسرین تلخی<sup>۲</sup>، مهدی جبباری نوقابی<sup>۲</sup>

<sup>۱</sup> دانشگاه فردوسی مشهد- دانشکده ریاضی- گروه آمار

<sup>۲</sup> دانشگاه علوم پزشکی مشهد- دانشکده بهداشت- گروه اپیدمیولوژی و آمار زیستی

چکیده: با توجه به پیشرفت روزافزون علوم مختلف به راحتی می‌توان به اهمیت کاربرد علم آمار در بسیاری از علوم به خصوص پزشکی پی برد. با توجه به ماهیت علوم پزشکی، مطالعات پزشکی از اهمیت و کاربرد بالایی برخوردار هستند. پس سازمان‌ها و افرادی که نتایج این تحقیقات را به کار می‌برند باید از صحت و دقت این نتایج اطمینان داشته باشند، این امر مستلزم این است که از همان ابتدا علم آمار به درستی و کامل برای دانشجویان این حیطه آموزش داده شود، تا از به وجود آمدن نتایج غلط و نادرست که ممکن است به کار گرفته شوند جلوگیری گردد. جدا از آموزش آمار، پژوهشگرانی نیز هستند که این علم را بنا به هر دلیلی اشتباه به کار برده و تفسیرهای نادرستی از تحقیقات خود ارائه می‌دهند. این اشتباهات نیز باید شناسایی شده و آموزش‌های لازم جهت رفع آنها ارائه گردد. واژه‌های کلیدی: آموزش آمار، روش‌های آماری، تحقیقات پزشکی، نتایج آماری، اشتباه‌های آماری، منابع خطا  
کد موضوع بندی ریاضی (۲۰۱۰): 07 – 62، 97B10، 97C70، 62Pxx.

### ۱ مقدمه

آمار، به عنوان یک رشته، بیش از ۳۰۰ سال سابقه دارد. در قرن بیستم، آمار به طور گسترده‌ای در سم‌شناسی، زیست‌شناسی مولکولی، آزمایش‌های بالینی و سایر زمینه‌های زیست‌پزشکی به کار گرفته شده است. دانشجویان رشته‌های علوم پزشکی و سایر علوم باید در طول دوره تحصیل، دانش و مهارت لازم را در زمینه یادگیری آمار و به کار گرفتن صحیح آن در تحقیقات حوزه خود کسب نمایند. آموزش محققان در استفاده از این علم امری بسیار مهم است که باعث بهبود کیفیت مطالعات در حوزه پزشکی می‌گردد. به صورت کلی، علوم تجربی وابستگی بسیار زیادی به تجزیه و تحلیل آماری به خصوص استنباط آماری دارند. با توجه به اینکه دست‌یابی به درک و منطق استنباط آماری در داده‌ها دشوار است، بنابراین تنها استفاده و تفسیر آن کافی نیست. استفاده نادرست از مفاهیم و روش‌های آماری باعث از دست رفتن زمان و هزینه می‌گردد و از همه مهم‌تر، از نظر اخلاق علمی، به علم و بشریت آسیب می‌رساند. از طرف دیگر، این موضوع می‌تواند باعث گمراه شدن سایر محققان و خدشه به تحقیقات علمی شود. استفاده نامناسب از آمار و روش‌های آن می‌تواند در هر مرحله از تحقیقات پزشکی مربوط به تجزیه و تحلیل داده‌ها، از جمله طراحی آزمایش، جمع‌آوری و پیش پردازش داده‌ها، تحلیل و تفسیر

نتایج مشاهده گردد. تقریباً ۵۰٪ از متون پزشکی دارای نقص آماری هستند و این اشتباهات و نقایص در مطالعات، بیشتر به دلیل عدم دانش آماری پژوهشگران رخ می‌دهد. با توجه به اینکه برآوردهای کارا، سازگار و ناریب با استفاده از علم آمار حاصل می‌شود، می‌توان با استفاده از آموزش آمار و اشاره به اشتباهات ناشی از نداشتن سواد آماری لازم، آموزش‌های مربوط به اجرای مطالعات علمی را از ابتدا تا پایان آن فراگرفت و به یک مطالعه با نتایج معتبر و دقیق دست یافت. تحلیل‌هایی که پژوهشگران در حوزه‌های علوم یا مهندسی کامپیوتر در زیرمجموعه علم داده (*DataScience*) یا داده‌کاوی (*DataMining*) انجام می‌دهند، مبنای آماری نداشته، در واقع از روش‌ها و مدل‌های آماری مختلف بدون رعایت پیش‌فرض‌های مربوطه استفاده می‌شود. پردازش داده‌های بزرگ (*BigData*) مانند داده‌ها با اندازه کوچک نیازمند استفاده از روش‌های آماری هستند. اینکه داده‌های بزرگ چطور تعریف شده‌اند مهم نیست، بلکه کلید یافتن اطلاعات ارزشمند از داده‌های بزرگ، تجزیه و تحلیل درست آماری داده‌ها است. در عصر داده‌های بزرگ نه تنها جمع‌آوری داده‌ها پویا است، بلکه فناوری ذخیره داده‌ها و فناوری پردازش داده‌ها نیز در هر زمان به روز می‌شوند، به این معنی که ابزارهای آماری پردازش داده‌ها نیز تغییر خواهند کرد. بر اساس تقاضای واقعی، آمار به تدریج همراه با تغییر تقاضا و داده‌ها توسعه می‌یابد. عصر داده‌های بزرگ بدون شک تقاضای زیادی برای تجزیه و تحلیل آماری ایجاد می‌کند. در طی چند دهه گذشته در مورد اصلاح آموزش آمار بحث‌هایی مطرح شده است. بسیاری از اساتید علاقه‌مند هستند تا به این سوال پاسخ بدهند، “چگونه می‌توان یادگیری دانشجویان را در آموزش آمار بهبود داد؟” اگر چه آمار مدرن دیدگاه‌های کاربردی و تقاضای زیاد در اشتغال ایجاد کرده است، اما دانشجویان همچنان از یادگیری آمار فاصله می‌گیرند. از آنجا که بیشتر تئوری‌ها در آمار مبتنی بر ریاضیات است، آموزش آمار را از سایر آموزش‌ها دشوارتر می‌سازد. بنابراین اساتید باید به راه‌حل‌های بهتری برای برطرف کردن این مشکلات در آموزش و تدریس دست یابند و به سؤال متداول، “چرا دانشجویان باید تئوری‌های آماری را از طریق ریاضیات یاد بگیرند اگر قرار است آمار سرشار از برنامه‌های کاربردی باشد” پاسخ دهند. در این مقاله ابتدا برخی از توصیه‌ها و روش‌های آموزش و یادگیری آمار که در انواع کلاس‌های آماری موفق به کار رفته معرفی خواهند شد. این روش‌ها در تدریس و یادگیری آمار می‌توانند باعث ارتقاء سواد آماری و ایجاد انگیزه در یادگیری آمار توسط دانشجویان گردند. پس از این، به چندین منبع خطا که محققان با نداشتن دانش و سواد آماری لازم، به راحتی در مطالعات پزشکی خود مرتکب می‌شوند، اشاره خواهد شد.

## ۲ توصیه‌ها و تکنیک‌های یادگیری و آموزش آمار

درک مفهومی از آمار به طور طبیعی حاصل نمی‌شود و به یک آموزش موفق مبتنی بر نوآوری و روش‌های مدرن، نیاز خواهد بود. مطالعاتی که بر تجدید نظر در مورد روش‌های سنتی تدریس آمار متمرکز شده‌اند، حاکی از تغییراتی هستند که باید انجام شود تا دانشجویان بتوانند آموزش‌هایی را دریافت کنند که هم به روز باشد و هم مرتبط با نیازهای جامعه باشد. برای پرداختن به یک سؤال مورد علاقه، مانند یک سؤال تحقیق در یک مطالعه پزشکی، نیاز به دانش در زمینه‌ی آن سؤال و درک مفهومی از آمار وجود دارد. در ادامه به معرفی برخی راهکارها پرداخته خواهد شد.

### ۱.۲ یادگیری ساختار یافته مطالب

تحقیقات در اوایل و اواسط دهه ۲۰۰۰ نشان داد فعالیت‌های یادگیری باید روی یادگیری ساختار یافته مطالب متمرکز شود. این امر به دانشجویان در دانش‌گزاره‌نویسی کمک می‌کند و درک مفهومی از آمار ایجاد می‌کند. برای دانشجویانی که تازه با مفاهیم آماری آشنا شده‌اند، ارائه مثال‌های کار شده به صورت جزئی یا کامل که به صورت فردی یا گروهی انجام شود، مورد نیاز می‌باشد. به عنوان مثال، به هنگام

معرفی یک موضوع آماری جدید به دانشجویان، بهتر است با مثال‌های کار شده یا کارهای تکمیل شده و یا با تمرین‌های رایانه‌ای ساده شروع شود و با پیشرفت دانشجویان در راستای حل مسئله و مشکلات عادی، راهنمایی‌ها کاهش یابد و حل دیگر مسائل به دانشجویان واگذار شود.

## ۲.۲ ارائه مثال‌های واقعی در رابطه با مطالب نظری

آموزش موثر با یک برنامه‌ریزی مناسب شروع می‌شود. دوره‌های درسی باید برنامه‌ریزی شود، فصل‌ها، واحدها و همچنین هر کلاس باید با مطالب و مثال‌هایی که با دقت و مختص همان کلاس انتخاب شده‌اند برنامه‌ریزی شود. برای مثال ممکن است یک موضوع پیچیده مانند آزمون فرضیه با مشکلی از یک خبر روزنامه امروز یا از مقاله اخیر آغاز شود. فرض کنید دولت ادعا می‌کند که نرخ بیکاری ۷٪ است، در حالی که احزاب مخالف ادعا می‌کنند که دولت این رقم را کم برآورد کرده است تا در انتخابات بعدی پیروز شود. واضح است که بین این دو گروه اختلاف نظر وجود دارد. چگونه می‌توان این اختلاف را با استفاده از آمار حل کرد؟ این مثال برای معرفی فرضیه صفر و فرضیه یک، از بیان تئوری آن‌ها خیلی مناسب‌تر است. بسیاری از مباحث در آمار می‌توانند با مقدمه‌ای پر جنب و جوش مانند مثال فوق آغاز شوند و اهداف را به درستی بیان کنند، پس از آن تشریح تئوری‌ها ارائه گردد. از این رو برنامه‌ریزی به عنوان اصلی‌ترین روند تدریس است.

## ۳.۲ نظریه بار شناختی

یکی از دیدگاه‌های یادگیری که حمایت فزاینده‌ای در زمینه آموزش پزشکی داشته است، نظریه بار شناختی است. در این نظریه، یادگیری با پیشرفت تدریجی از طرح‌واره‌های شناختی دانشجویان صورت می‌گیرد. هنگامی که دانشجویان پزشکی برای اولین بار با آمار مواجه می‌شوند، آن‌ها به طور معمول دارای طرح‌واره‌های شناختی ناقصی در مورد موضوع مورد نظر هستند. توضیح دادن به خود، استدلال و تشریح مطالب توسط یک همکار برای همکار دیگر فرایندهایی هستند که به یادگیری کمک کرده و در صورت دریافت میزان مناسبی آموزش، به کامل شدن و تفهیم موضوع می‌انجامد.

## ۴.۲ فراهم آوردن منابع یکپارچه

موضوع دیگری که می‌تواند در یادگیری موثر باشد، فراهم آوردن یا معرفی یک منبع یکپارچه از مطالب به دانشجویان می‌باشد. به عنوان مثال فرض کنید دانشجویان برای درک یک مطلب یا حل یک مشکل، بین صفحات وب جلو و عقب بروند، منابع کافی برای موضوع موجود نباشد و یا با منابع نامعتبر رو به رو شوند، این منجر به دور شدن از موضوع و عدم یادگیری خواهد شد. بنابراین نیاز به ارائه منابع اختصاصی برای مقابله با دریافت نامطلوب اطلاعات می‌باشد. به عنوان راه‌حلی دیگر، شرکت دانشجویان در یک کارگاه عملی و دریافت توضیحات لفظی در مورد یک مبحث، چگونگی استفاده از یک ابزار یا نرم افزار آماری خاص و... می‌تواند موثر باشد و به یکپارچه‌سازی منابع موجود کمک نماید.

## ۵.۲ آموزش تصویری برخی مفاهیم

برخی مفاهیم باید به صورت بصری و تصویری توضیح داده شوند تا آموزش موثر و یادگیری انجام شود. مثلاً برای معرفی توزیع نرمال، کشیدن منحنی نرمال یا همان شکل زنگوله‌ای به درک دانشجویان در شناخت توزیع نرمال کمک بسیاری می‌کند. با ظهور یوتیوب و سایر رسانه‌ها، تعداد فیلم‌هایی که مفاهیم خاص را در بستر زندگی واقعی شبیه‌سازی می‌کنند، بصورت تصاعدی در حال افزایش است. البته، نمی‌توان انتظار داشت دانشجویانی که هنوز با مفاهیم آشنا نیستند، یاد بگیرند که کدام فیلم‌ها دارای کیفیت خوب و کدام یک خوب نیست (مطالب کاملاً آنلاین وجود دارند که از کیفیت مطلوبی برخوردار نیستند).

## ۶.۲ بیان روشن از محتوای مطالب

مطالعات مختلفی از این فرض تاثیر می‌پذیرد که اثربخشی آموزش تا حد قابل توجهی به معیارهای ارزیابی در یک دوره یا برنامه درسی بستگی دارد. معیارهای ارزیابی و فعالیت‌های یادگیری چالش برانگیزتر ممکن است یادگیری را تحریک کنند. اساتید و طراحان آموزشی نیاز به یکپارچه نمودن موضوع دارند. به عنوان مثال، باید آمار در چارچوب برنامه درسی پزشکی قرار داده شود و تصمیم گرفته شود که در چه دوره‌ای و با چه مطالب و فعالیت‌هایی، یادگیری پوشش داده شود. برای کامل کردن موضوع آمار در برنامه درسی، باید پایان روشنی در مورد آنچه دانشجویان انتظار دارند در پایان برنامه درسی و پس از گذراندن یک دوره خاص در آن برنامه درسی کسب کنند، داشته باشند. به عبارتی دیگر، در ابتدای هر مبحث، کاربرد آن با مثال‌های واقعی توضیح داده شود.

## ۷.۲ آموزش سلسله مراتبی مفاهیم

مفاهیم آماری باید با سلسله مراتب ارائه شده و از مفاهیم ساده به سمت مفاهیم پیچیده‌تر حرکت کرد. به عنوان مثال، نمی‌توان انتظار داشت دانشجویان درک صحیحی از ضریب همبستگی پیرسون بدون درک مفاهیمی که بر اساس آن پایه‌گذاری شده‌اند (مانند میانگین حسابی، انحراف استاندارد، کواریانس و استاندارد سازی) داشته باشند. به عبارت دیگر، تا زمانی که دانشجویان این مفاهیم اساسی‌تر را درک نکرده باشند، باید قبل از یادگیری ضریب همبستگی پیرسون، حول این مفاهیم بچرخند. یا مثلاً در رگرسیون خطی ساده و چندگانه ابتدا باید همه پذیره‌های زیربنایی، انواع همبستگی، انواع ضرایب رگرسیون و موارد دیگر به درستی درک شوند.

## ۸.۲ دوره‌های آموزشی آنلاین

بسیاری از دوره‌های آموزشی با کیفیت بالا که در دانشگاه‌های درجه یک جهان ارائه می‌شود بصورت آنلاین و رایگان در اختیار کاربران قرار داده می‌شود، به طوری که منابع آموزشی با کیفیت بالا با کمترین هزینه برای هر فردی در جهان که مایل به پذیرش و یادگیری است فراهم می‌شود. این نوع آموزش و ایده‌های توسعه آموزش مبتنی بر داده‌های بزرگ، تأثیر زیادی بر آموزش سنتی خواهد گذاشت. کلمه *MOOC* از حروف اول کلمات عبارت *Massive Open Online Course* گرفته شده و به معنای دوره آنلاین باز و بزرگ است. ظهور *MOOC* و سایر سیستم‌های آموزش آنلاین مانند *edX*, *Udacity*, *Coursera* در زمینه‌های تدریس داخلی و خارجی به موضوعی داغ تبدیل شده است و باعث ایجاد یک انقلاب در یادگیری و تعلیم و تربیت می‌شود که در حال تغییر در شیوه تدریس سنتی است. در عصر داده‌های بزرگ، کسب دانش دیگر محدود به کلاس نخواهد بود و آموزش آماری نیز باید به روز بوده و به دنبال پیشرفت در اصلاحات باشد. این نوع شیوه تدریس برای دانشجویان مناسب است تا مطالبی را که قبلاً آموخته‌اند بررسی و درک کنند.

## ۹.۲ استفاده کمتر از فرمول های ریاضی

از آنجا که هدف از آموزش آمار در یک زمینه پزشکی، کمک به دانشجویان است تا درک مفهومی خود را از آمار توسعه دهند، بیان بیش از اندازه آنها ممکن است دانشجویان را از اصل میحث و هدف دور نگه دارد. بنابراین بیان به اندازه تئوری این علم برای تحقیقات آنها کافی است. به عنوان مثال بیان فرمول های اساسی برای انحراف استاندارد یا میانگین حسابی ممکن است درک فرد از این مفاهیم و رابطه آنها با هم را تسهیل کند اما بیان فرمول ها و ریاضیات پیشرفته تر ممکن است باعث منحرف شدن از هدف واقعی یعنی فهم مفهومی گردد.

## ۱۰.۲ واگذاری پروژه به دانشجویان

یک راه آسان برای آموزش صحیح آمار در رشته های دیگر، واگذار کردن تحقیقات و پژوهش های کوچک مربوط به گرایش در دوره های آموزش (به عنوان مثال در پزشکی تحقیقات در مورد آناتومی، قلب و عروق و...) است. به این ترتیب دانشجویان می توانند با تجزیه و تحلیل و گزارش اطلاعات آماری در زمینه ای که برای آنها آشنا و جالب است، دانش خود را در مورد یک موضوع به دست آورند. علاوه بر این، این رویکرد می تواند دانشجویان را در توانایی نوشتن پایان نامه یا مقالات بزرگتر که اطلاعات آماری را ارائه می کند، کمک کند.

## ۱۱.۲ مورد بحث قرار دادن تصورات اشتباه

بهتر است تصورات اشتباه در ذهن دانشجویان مطرح و به بحث گذاشته و سپس نتیجه واحد گرفته شود. برای رفع این تصورات می توان جلسات پرسش و پاسخ برگزار کرد. البته محیط حاکم باید به گونه ای طراحی شود که دانشجویان تشویق به این کار شده و در مورد برداشت های خود از مطالب و مفاهیم را به بحث و گفتگو بپردازند. این امر منجر به رفع برخی ابهامات در مورد مفاهیم آماری از ذهن دانشجویان خواهد شد تا مطالب بعدی راحت تر درک شوند.

## ۱۲.۲ مشارکت و برقراری ارتباط با پژوهشگران رشته آمار

جهت یادگیری بهتر، دانشجویان باید با یکدیگر برای حل مسائل مشورت نمایند و در به اشتراک گذاری ایده ها همکاری کنند. در یادگیری مطالب و کار با نرم افزارهای آماری، تأمل و کاوش در مفاهیم آماری و آنالیز داده ها و حل مشکلات دانشجویان به صورت مشترک و گروهی با یکدیگر همکاری کنند. اساتید فعالیت هایی که دانشجویان را قادر به ایجاد دانش و یادگیری می کنند، معرفی نمایند. در این میان اگر این ارتباطات گسترش یافته و دانشجویان با پژوهشگران و آمارشناسان در تبادل اطلاعات قرار گیرند تا با مشورت و توضیحات این پژوهشگران درک مفاهیم و تحلیل های آماری برای دانشجویان به صورت صحیح و آسان انجام گیرد. این ارتباطات می تواند با شرکت کردن دانشجویان در سخنرانی های آمار مقدماتی، سمینارهای تحقیقاتی و... انجام پذیرد.

## ۱۳.۲ استفاده از تکنولوژی در یادگیری

استفاده از تکنولوژی در آموزش آمار باعث تسریع در آموزش مطالب شده و به نوعی یادگیری را بهبود می بخشد. در زیر به چند مورد از مزایای استفاده از تکنولوژی را بیان می کنیم. ۱- کاهش تمرکز روی محاسبات و استفاده از کامپیوتر و ماشین حساب برای انجام محاسبات باعث می شود در زمان اندک و با دقت بالایی محاسبات انجام شود و دانشجویان وقت بیشتری برای تمرکز بر روی درک مفاهیم صرف کنند. ۲- استفاده از ماکروهای موجود و توجه کمتر روی ورود اطلاعات، دستکاری اعداد در فرمول ها و همچنین استفاده از

تمرین‌هایی که فقط از مجموعه داده‌های کوچک و آموزشی استفاده می‌کنند، باعث صرفه جویی در زمان می‌گردد. برای ترغیب دانشجویان به بررسی عمیق‌تر داده‌ها، باید از فناوری و تکنولوژی در کلاس استفاده کرد، تا داده‌ها بتوانند نمایشی از وضعیت و هدف موجود را به دانشجویان نشان دهند تا دانشجویان بتوانند موضوعات مفهومی مرتبط را مورد توجه قرار دهند. ۳- دسترسی آسان به مشاهده و طراحی شبیه سازی‌ها به عنوان یک ابزار آموزشی می‌تواند نقش مهمی در تقویت توانایی دانشجویان در مطالعه فرآیندهای تصادفی و مفاهیم آماری داشته باشد. با چنین شبیه سازی‌هایی، مفاهیم انتزاعی مانند توزیع نمونه برداری و فاصله اطمینان در ذهن تثبیت خواهند شد. درک دانشجویان با انجام این تکرارها، کنترل پارامترها (اندازه نمونه و تعداد تکرارها) و رفتاری که مشاهده می‌کنند در مقایسه با تکیه بر صرفاً مطالب نظری، بسیار بالا خواهد بود. ۴- آموزش کار با نرم افزارهای آماری، برنامه‌های شبیه سازی شده و استفاده از منابع موجود بصورت آنلاین در کلاس، به دانشجویان امکان می‌دهد مفاهیم آماری را کشف و بهتر درک کنند.

### ۳ منابع اصلی خطا در تحقیقات

همانطور که اشاره شد در این بخش به یک سری از اشتباه‌ها و خطاهای رایج آماری در تحقیقات که بیشتر پژوهشگران در زمینه علم پزشکی آنها را مرتکب می‌شوند اشاره خواهیم کرد و برای عدم تکرار این اشتباه‌ها راهکارهایی پیش روی پژوهشگران قرار خواهیم داد.

#### ۱.۳ توصیف جامعه

جامعه آماری یکی از مباحث اولیه در تحقیق به شمار می‌آید که تعیین صحیح آن برای حصول نتایج مورد قبول بسیار حائز اهمیت می‌باشد. جامعه مورد مطالعه محقق باید از نظر زمان، مکان و حداقل یک ویژگی مشترک خاص در واحدهای آن تعریف شده باشد. در این صورت انتخاب واحدهای آماری و تعیین متغیرهایی که مورد تجزیه و تحلیل قرار خواهند گرفت، به آسانی صورت خواهد پذیرفت. در غیر این صورت محققان در انتخاب واحدهای نمونه دچار مشکل خواهند شد و منجر به افزایش ناهمگونی واحدهای آماری خواهد شد. مزیت خوب دیگر تعریف دقیق جامعه آماری به وضوح مشخص شدن، متغیرهای مورد مطالعه است.

#### ۲.۳ طرح نمونه‌گیری

در تحقیقات، روش نمونه‌گیری متناسب با نوع پژوهش انتخاب می‌شود که در آن باید موضوع مطالعه، خصوصیات جامعه، دوره تحقیق و هزینه تهیه مشاهدات در نظر گرفته شود، که به نوعی اهمیت نمونه‌گیری را بیان می‌کند. بیش از یک روش نمونه‌گیری می‌تواند در یک مطالعه مورد استفاده قرار گیرد. یکی از اشتباه‌های رایجی که محققان مرتکب می‌شوند، استفاده از همان روش نمونه‌گیری است که در مطالعات مشابه استفاده شده است. اگر این روش نمونه‌گیری برای داده‌های تحقیق حاضر مناسب نباشد، در این صورت نتایج حاصل از نمونه دارای اریبی خواهد بود و غیر مستقیم منجر به تفسیرهای نادرست خواهد شد. موضوع بعدی که حائز اهمیت است، تعیین و تعریف واضح ملاک‌های ورود و خروج واحدهای آماری از نمونه است. از رایج‌ترین اشتباه‌ها در انتخاب واحدها، جمع آوری واحدها توسط محققانی است که آموزش لازم را ندیده‌اند، بنابراین در این صورت ما نمونه‌ای خواهیم داشت که ویژگی جامعه را به خوبی نشان نداده و گویای جامعه نیست. از این رو معیارهای انتخاب نمونه باید به طور واضح بیان شوند. نمونه‌گیری‌های تصادفی و غیرتصادفی دو نوع نمونه‌گیری است که در تحقیقات مورد استفاده قرار می‌گیرند. پس باید متناسب با نوع پژوهش و جامع مورد تحقیق مورد بررسی قرار گرفته و یکی را انتخاب کرد. که در بیشتر تحقیق‌ها مشاهده شده که پژوهشگران در انتخاب آن دچار اشتباه شده‌اند. با افزایش تعداد

افراد در نمونه، نمونه نماینده بهتری از جامعه خواهد بود و وضعیت جامعه را بهتر نشان خواهد داد. حجم نمونه مناسب برای مطالعات باید با توجه به مطالعات قبلی، با یک خطای معین و سطح معنی داری و توان آزمون صورت پذیرد. اما برخی از محققان فقط با داشتن اطلاعاتی مانند میانگین نسبت، انحراف معیار، خطای معیار میانگین و ... حجم نمونه تعیین کرده و به منابع ارجاع می‌دهند. نکته دیگر در تعیین حجم نمونه این است که در بعضی موارد محققان برای دست یابی هر چه زودتر به نتایج حجم نمونه کمتری در نظر می‌گیرند. پس با توجه به مطالب ارائه شده در طرح نمونه‌گیری باید روش نمونه‌گیری، معیارهای انتخاب نمونه، نوع نمونه‌گیری و تعیین حجم نمونه باید با دقت بیشتری بررسی شوند چون در این موارد احتمال خطا توسط محققین بسیار زیاد است.

### ۳.۳ طراحی مطالعه و کسب اطلاعات در مورد متغیرها

با توجه به اهمیت طرح مطالعه، اگر پژوهشگران طرح مطالعه را نادرست انتخاب کنند، نتایج با دقت پایین برآورد خواهند شد. کارآزمایی‌های بالینی تصادفی شده دارای گروه کنترل قدرتمندترین طرح ممکن اما هزینه‌بر در تحقیقات پزشکی هستند. مطالعات مشاهده‌ای وقت‌گیر اما بسیار سریع و ارزان هستند. با توجه به اینکه نتایج مطالعات مقطعی مختص همان زمان هستند لذا نباید از نتایج آن برای تصمیم‌گیری در مقاطع زمانی دیگر استفاده کرد. قبل از شروع تحقیق باید متغیرهایی که در پژوهش‌های قبلی مورد مطالعه قرار گرفته‌اند شناسایی شوند تا از در ارتباط بودن با موارد پژوهش اصلی اجتناب گردد، مگر اینکه یک رابطه علت و معلولی بین آن متغیرها وجود داشته باشد. به عنوان مثال در بررسی رابطه بین مصرف الکل و سرطان ریه، باید عادت به سیگار کشیدن که معمولاً با مصرف الکل استفاده می‌شود، به عنوان یک عامل مخدوش کننده مورد استفاده قرار بگیرد.

### ۴.۳ جمع آوری و پیش پردازش داده‌ها

در برخی پژوهش‌ها ممکن است متغیرها با ابزارهای نامناسبی اندازه‌گیری گردند که در این صورت نتایجی که از تحلیل این اطلاعات به دست می‌آید گمراه کننده بوده و یا ناخوشایند خواهد بود. حتی ممکن است باعث رد یا پذیرش نا به جای ادعاهای محقق گردد. به عنوان مثال برخی از پژوهشگران در حین بررسی تأثیر مصرف سیگار بر روی یک بیماری، افراد را به سیگاری و غیر سیگاری طبقه بندی می‌کنند. در این حالت مواردی مانند اینکه این موضوع چقدر طول کشیده یا اینکه فرد در روز چند عدد سیگار می‌کشد مشاهده نمی‌شود. برای اینکه مشاهدات حاوی اطلاعات دقیقتر و بیشتری باشند متغیرها به صورت بسته - سال برای اندازه‌گیری میزان سیگار کشیدن در نظر گرفته می‌شود (به عنوان مثال برای شخصی که ۶ سال و ۱۰ سیگار در روز کشیده است مقدار مشاهده  $3 = \frac{6 \times 10}{3}$  خواهد بود). در مرحله گرد آوری داده‌ها، اول باید منبع داده‌ها مشخص شوند، پس از مشخص شدن منبع باید به طور کامل و دقیق از نظر قابل اعتماد بودن بررسی شود. در این مرحله یکی از خطاهای رایج، استفاده از داده‌های ثبتی‌ای که توسط گروه دیگری گردآوری شده. تعریف قلم و متغیرهای آماری، اگر به صورت جامع و کامل با یک شناسنامه معین نباشد، باعث برداشت‌های متفاوت از آن شاخص با قلم آماری شده است، و لذا داده‌های ثبت شده توسط افراد مختلف، دامنه متفاوت داشته، بنابراین دارای صحت و دقت لازم نمی‌باشند. در مواردی ممکن است محققان متغیر دقیق مورد نظر خود را در تحقیق‌های گذشته پیدا نکنند یا ممکن است آنهار در مقیاس‌های متفاوتی سنجیده شده باشند. یکی دیگر از متداول‌ترین منابع خطا در مطالعات این است که برخی از آزمودنی‌ها قبل از پایان مطالعه پیشامد را کسب نکرده یا فرد در طول مدت مطالعه گم‌شده یا از دست رفته باشد و یا فرد به علت دیگری جز پیشامد مورد نظر از مطالعه خارج شود و یا در مرحله جمع آوری، اطلاعاتی در مورد آنها کسب نشود. اگر در مجموعه داده‌های تحقیق چنین داده‌هایی وجود داشت که به اصطلاح داده‌های سانسور شده نامیده می‌شوند، باید اطلاعاتی در مورد آنها به دست آورد یا اگر ارزیابی شده باشند باید قید شود که در چه مرحله از تحقیق



حذف شده اند. کنار گذاشتن این موارد باعث کاهش قدرت مطالعه در شروع آن تحقیق می شود. با توجه به اینکه داده‌ها از نظر مقیاس به چهار دسته نسبتی، فاصله‌ای، ترتیبی و اسمی تقسیم بندی می‌شوند، پژوهشگران غیرآماري تمایل دارند از دو مقیاس فاصله‌ای و نسبتی به دلیل ویژگی‌های ریاضی شان استفاده نمایند. برخی از پژوهشگران با اینکه متغیرها دارای مقیاس نسبتی یا فاصله‌ای هستند آن‌ها را به مقیاس‌های اسمی تبدیل و به تجزیه و تحلیل می‌پردازند، که این تغییر در نوع اندازه‌گیری باعث کاهش سطح دقت در اندازه‌گیری متغیرها شده و باعث از دست رفتن اطلاعات و نهایتاً منجر به تفسیر نادرست از پژوهش خواهد شد. به عنوان مثال، به جای مقایسه مقادیر کلسترول به دست آمده در قبل و بعد از استفاده از دارو برای کاهش کلسترول، اگر داده‌ها به صورت پایین - متوسط - بالا طبقه بندی شوند باعث از دست رفتن اطلاعات و نادیده گرفتن تغییرات کلسترول خواهد شد و به دنبال آن نتیجه‌گیری‌های نادرست حاصل خواهد شد.

### ۵.۳ اجرا و روش تجزیه و تحلیل

از خطاهایی که محققان با آن روبه‌رو هستند این است که اگر پژوهشی مشابه یافت شود، در سدد استفاده از همان روش تحلیل آماری، روشها و آزمون‌هایی را در پژوهش خود دارند که در موارد یافت شده، استفاده شده است. روش آماری، مناسب برای پژوهش و مجموعه داده است که با معیارهای آماری مانند تعداد داده‌ها، مقیاس داده‌ها، تغییرپذیری و توزیع نظری متناسب باشد. همچنین انتخاب نادرست نرم‌افزارهای آماری و روش محاسبه آماری زمانی اتفاق می‌افتد که محققان از اشخاص متخصص آماری مشاوره نمی‌گیرند و بر پایه دانش اندک خود از این علم و یا به روش‌های موجود در متون گذشته اکتفا می‌کنند. که این خطا نیز در بسیاری از مواقع باعث تحلیل نادرست از پژوهش می‌شود. گاهی اوقات محققان از روش‌های آماری که برای آن‌ها آسانتر بوده استفاده می‌کنند و با به دست آوردن مقدار P- تصور می‌کنند که محاسبات آنها درست بوده و نظرها و تفسیرهای خود را اعمال می‌کنند. برای مثال برای رگرسیون نمایی در تجزیه و تحلیل بقا برخی از نرم‌افزارها از مدل مخاطره متناسب به صورت  $\lambda(\frac{t}{z}) = \lambda e^{\beta z}$  نمایش داده و به کار می‌گیرند و بعضی از مدل لوگ - خطی  $(\log(T) = -\alpha + \beta^* z)$  استفاده می‌کنند. که نتایج آنها با علامت‌های مخالف هم مشخص می‌شوند  $(\hat{\beta}^* = \hat{\beta})$ . بنابراین در بعضی مواقع نتایج به دلیل متفاوت بودن نتایج و نحوه محاسبات نرم‌افزارها و برنامه‌های مختلف باید توسط پژوهشگران بازنگری شود. در بعضی مواقع پژوهشگران از متفاوت بودن نتایج خود با نتایج مطالعات مشابه نگران بوده و تصور می‌کنند علت این ایراد ناکافی بودن تعداد نمونه بود و ممکن است تعداد افراد را تا رسیدن به نتیجه مورد نظر افزایش دهند و یا برخی از آنهایی را که نتایج متفاوت دارند از نمونه خارج کنند. و یک اشتباه دیگری که ممکن است مرتکب شوند این است که محققان نوع آزمون را به دلیل انتظارشان انتخاب می‌کنند. در صورت قبول فرضیه صفر محققان در تلاش این هستند که از آزمون دیگری استفاده کنند.

### ۶.۳ تفسیر

در برخی از پژوهش‌ها مشاهده می‌شود، پژوهشگران به دلیل اینکه نمی‌دانند چه برداشت و نتیجه‌ای از اعداد می‌توان داشت یا چون ادبیات لازم را برای تفسیر ندارند، اعداد به دست آمده را بدون تفسیر و معانی فقط ذکر می‌کنند. بنابراین عبارات گمراه کننده آماری را به کار می‌برند. به همین دلیل قبل از انتشار مطالعات باید نتایج توسط یک آمارشناس بازبینی گردد. به عنوان مثال بسیاری از پژوهشگران تفاوت بین انحراف استاندارد و خطای استاندارد را نمی‌دانند و اشتباه گزارش می‌کنند. در مقایسه گروه‌های وابسته، اگر برای یک یا چند متغیر، مشاهده گمشده وجود داشته باشد، برخی از پژوهشگران این موضوع را که بعضی از نرم‌افزارهای آماری تجزیه و تحلیل را با استفاده از داده‌های گمشده انجام می‌دهند، درک نکرده و در تفسیر گروه‌های وابسته آمار توصیفی کل مجموعه داده‌ها را در نظر می‌گیرند.

برخی از محققان با به دست آوردن نتایج معنادار، از عبارت‌هایی چون ” نتایج از نظر آماری معنادار نیست اما میانگین  $x$  بزرگتر از  $y$  است ” استفاده می‌کنند، این عبارت با نتیجه آزمون متناقض هست در صورتی که اگر مطالعه با شرایط و نمونه یکسان تکرار شود ممکن است میانگین  $x$  کوچکتر شود اما – مقدار بزرگتر از  $0.05$  حاصل شود.

## ۴ بحث و نتیجه‌گیری

در پژوهش حاضر برخی راهکارهای موثر بر یادگیری و آموزش آمار و همچنین برخی اشتباهات رایج در سایر رشته‌ها، به‌خصوص علوم پزشکی با هدف ارتقای سواد آماری و کمتر شدن خطاها در تحقیقات پزشکی معرفی شدند. این راهکارها جهت آموزش آمار در برخی کلاس‌های موفق آماری به کار رفته‌اند و لذا به نظر می‌رسد این رویکردها با همکاری و تعامل بین اساتید و دانشجویان محقق خواهد شد. با دقت در مطالب ارائه شده، کادر آموزشی در دانشگاه‌ها باید هر برنامه آموزشی مناسب برای دانشجویان را ارائه دهند، به گونه‌ای که اساتید بتوانند بازخورد تدریس را به صورت مؤثرتر و عینی‌تری به دست آورند، کیفیت آموزشی را بهبود بخشند و استعدادها را بزرگتری را پرورش دهند. علم آمار باید بیشتر در خدمت و برآورده کردن تقاضای زمینه‌های مختلف به عنوان هدف باشد، پیشنهاد می‌شود که شیوه آموزش آمار در دانشگاه‌ها تغییر یابد و به دنبال پیشرفت در اصلاح روش‌های آموزشی باشند. از طرف دیگر، با توجه به اهمیت و حیاتی بودن نتایج تحقیقات در حوزه علوم پزشکی، دانشجویان باید علم آمار و مباحث آماری را جزو ضروری‌ترین بخش تحقیق دانسته و در تمام مراحل در نظر بگیرند. جزئیات و مفاهیم را به خوبی درک کرده و اشتباهاتی را که امکان دارد در طراحی مطالعه، جمع‌آوری و پیش‌پردازش داده‌ها، تجزیه و تحلیل و تفسیر، مرتکب شوند، به خوبی فراگیرند. با این حال، وجود یک کارشناس آمار در تیم تحقیقاتی آن‌ها ضروری است و باید در تمام مراحل تحقیق، حتی قبل از شروع تحقیق با آمارشناس مشورت شود تا نتایج حاصل از حداکثر دقت لازم برخوردار باشند.

## مراجع

- [1] Chance, B., Ben-Zvi, D., Garfield, J., Medina, E. (2007). Technology innovations in statistics education. *Technol. Innov. Stat. Educ*, 1(1).
- [2] Ercan, I., Yazici, B., Yang, Y., Ozkaya, G., Cangur, S., Ediz, B., Kan, I. (2007). Misusage of statistics in medical research. *Eur J Gen Med*, 4(3), 128-134.
- [3] Garfield, J. (1995). How students learn statistics. *International Statistical Review/Revue Internationale de Statistique*, 25-34.
- [4] Glencross, M. J., Binyavanga, K. W. (1996). The role of technology in statistics education: A view from a developing country. *Role of Technology*, 1, 301-307.
- [5] Higazi, S. M. (2002). Teaching statistics using technology. *ICOTS6 Proceedings*.

- [6] Leppink, J. (2017). Helping medical students in their study of statistics: A flexible approach. *Journal of Taibah University Medical Sciences*, 12(1), 1-7.
- [7] Li, Y., Huang, C., Zhou, L. (2018). Impacts on Statistics Education in Big Data Era. *Kuram ve Uygulamada Egitim Bilimleri*, 18(5), 1236-1245.
- [8] Peiris, M. S. (2012). A way of teaching statistics: An approach to flexible learning. *International Journal of Innovation in Science and Mathematics Education (formerly CAL-laborate International)*, 9(1).
- [9] Strasak, A. M., Zaman, Q., Pfeiffer, K. P., Göbel, G., Ulmer, H. (2007). Statistical errors in medical research-a review of common pitfalls. *Swiss medical weekly*, 137(0304).
- [10] Tishkovskaya, S., Lancaster, G. A. (2010, July). Teaching strategies to promote statistical literacy: review and implementation. In *Data and Context in Statistics Education: Towards an Evidence-Based Society*. Proceedings of the Eighth International Conference on Teaching Statistics. Voorburg, The Netherlands: International Statistical Institute.