



دانشگاه سیام نور

آمار در شیمی تجزیه

دکتر ملیحه مجرب

دکتر عزیزاله نژادعلی

«امروزه کتاب‌خوانی و علم‌آموزی، نه تنها یک وظیفه‌ی ملی، که یک واجب دینی است.»^۱

در عصر حاضر یکی از شاخصه‌های ارزیابی رشد، توسعه و پیشرفت فرهنگی هر کشوری میزان تولید کتاب، مطالعه و کتاب‌خوانی مردم آن مرز و بوم است. ایران اسلامی نیز از دیرباز تاکنون با داشتن تمدنی چندهزارساله و مراکز متعدد علمی، فرهنگی، کتابخانه‌های معتبر، علما و دانشمندان بزرگ با آثار ارزشمند تاریخی، سرآمد دولت‌ها و ملت‌های دیگر بوده و در عرصه‌ی فرهنگ و تمدن جهانی به‌سان خورشیدی تابناک همچنان می‌درخشد و با فرزندان نیک‌نهاد خویش هنرنمایی می‌کند. چه کسی است که در دنیا با دانشمندان فرزانه و نام‌آور ایرانی همچون ابوعلی سینا، ابوریحان بیرونی، فارابی، خوارزمی و ... همچنین شاعران برجسته‌ای نظیر فردوسی، سعدی، مولوی، حافظ و ... آشنا نباشد و در مقابل عظمت آنها سر تعظیم فرود نیاورد. تمامی این افتخارات ارزشمند، برگرفته از میزان عشق و علاقه فراوان ملت ما به فراگیری علم و دانش از طریق خواندن و مطالعه منابع و کتاب‌های گوناگون است. به شکرانه‌ی الهی، تاریخ و گذشته ما، همیشه درخشان و پربار است. ولی اکنون در این زمینه در چه جایگاهی قرار داریم؟ آمار و ارقام ارائه‌شده از سوی مجامع و سازمان‌های فرهنگی در مورد سرانه‌ی مطالعه‌ی هر ایرانی، برآیمان چندان امیدوارکننده نمی‌باشد.

کتاب، دروازه‌ای به سوی گستره‌ی دانش و معرفت است و کتاب خوب، یکی از بهترین ابزارهای کمال بشری است. همه‌ی دستاوردهای بشر در سراسر عمر جهان، تا آنجا که قابل کتابت بوده است، در میان دست‌نوشته‌هایی است که انسان‌ها پدید آورده و می‌آورند. در این مجموعه‌ی بی‌نظیر، تعالیم الهی، درس‌های پیامبران به بشر، و همچنین علوم مختلفی است که سعادت بشر بدون آگاهی از آنها امکان‌پذیر نیست. کسی که با دنیای زیبا و زندگی‌بخش کتاب ارتباط ندارد بی‌شک از مهم‌ترین دستاورد انسانی و نیز از بیشترین معارف الهی و بشری محروم است. با این دیدگاه، به‌روشنی می‌توان ارزش و مفهوم رمزی عمیق در این حقیقت تاریخی را دریافت که اولین خطاب خداوند متعال به پیامبر گرامی اسلام(ص) این است که «بخوان!» و در اولین

۱. پیام مقام معظم رهبری به مناسبت آغاز هفته کتاب ۷۲/۱۰/۴

سوره‌ای که بر آن فرستاده‌ی عظیم‌الشأن خداوند، فرود آمده، نام «قلم» به تجلیل یاد شده‌است: «إِقْرَأْ وَ رَبُّكَ الْأَكْرَمُ. الَّذِي عَلَّمَ بِالْقَلَمِ» در اهمیت عنصر کتاب برای تکامل جامعه‌ی انسانی، همین بس که تمامی ادیان آسمانی و رجال بزرگ تاریخ بشری، از طریق کتاب جاودانه مانده‌اند.

دانشگاه پیام‌نور با گستره‌ی جغرافیایی ایران‌شمول خود با هدف آموزش برای همه، همه‌جا و همه‌وقت، به‌عنوان دانشگاهی کتاب‌محور در نظام آموزش عالی کشورمان، افتخار دارد جایگاه اندیشه‌سازی و خردورزی بخش عظیمی از جوانان جویای علم این مرز و بوم باشد. تلاش فراوانی در ایام طولانی فعالیت این دانشگاه انجام پذیرفته تا با بهره‌گیری از تجربه‌های گرانقدر استادان و صاحب‌نظران برجسته کشورمان، کتاب‌ها و منابع آموزشی درسی شاخص و خودآموز تولید شود. در آینده هم، این مهم با هدف ارتقای سطح علمی، روزآمدی و توجه بیشتر به نیازهای مخاطبان دانشگاه پیام‌نور با جدیت ادامه خواهد داشت. به‌طور قطع استفاده از نظرات استادان، صاحب‌نظران و دانشجویان محترم، ما را در انجام این وظیفه‌ی مهم و خطیر یاری‌رسان خواهد بود. پیشاپیش از تمامی عزیزانی که با نقد، تصحیح و پیشنهادهای خود ما را در انجام این وظیفه‌ی خطیر یاری می‌رسانند، سپاسگزاری می‌نماییم. لازم است از تمامی اندیشمندانی که تاکنون دانشگاه پیام‌نور را منزلگاه اندیشه‌سازی خود دانسته و ما را در تولید کتاب و محتوای آموزشی درسی یاری نموده‌اند، صمیمانه قدردانی گردد. موفقیت و بهروزی تمامی دانشجویان و دانش‌پژوهان عزیز آرزوی همیشگی ما است.

دانشگاه پیام‌نور

فهرست مطالب

پیشگفتار	نه
فصل اول: مفاهیم اولیه	۱
هدف کلی	۱
هدف‌های یادگیری	۱
مقدمه	۲
۱-۱ انواع خطا در نتایج و داده‌ها	۳
۱-۱-۱ خطاهای معین	۳
۲-۱-۱ خطاهای نامعین	۵
۲-۱ برخی اصطلاحات آماری	۵
۱-۲-۱ دقت و صحت داده‌ها	۵
۲-۲-۱ خطای مطلق	۷
۳-۲-۱ خطای نسبی	۷
۴-۲-۱ میانگین حسابی	۸
۵-۲-۱ میانه	۹
۶-۲-۱ مُد	۹
۷-۲-۱ دامنه یا گستره	۱۰
۸-۲-۱ انحراف از میانگین	۱۱
۹-۲-۱ انحراف استاندارد	۱۱
۱۰-۲-۱ واریانس	۱۴
۳-۱ توزیع نرمال	۱۵
۱-۳-۱ حدود اطمینان و فاصله اطمینان	۱۷
۴-۱ ارقام بامعنی و انتشار خطای محاسباتی	۲۲
۱-۴-۱ ارقام بامعنی در محاسبات	۲۳
۵-۱ انتشار خطا (عدم قطعیت) در محاسبات	۲۷
خلاصه فصل اول	۳۱

۳۳	خودآزمایی چهارگزینه‌ای فصل اول
۳۴	خودآزمایی تشریحی فصل اول
۳۷	فصل دوم: آزمون‌های معنی‌دار بودن
۳۷	هدف کلی
۳۷	هدف‌های یادگیری
۳۸	مقدمه
۳۸	۱-۲ آزمون مقایسه میانگین تجربی با مقدار معلوم
۴۰	۲-۲ آزمون t جفت‌شده
۴۲	۳-۲ آزمون F و مقایسه دقت اندازه‌گیری‌ها
۴۴	۴-۲ خطای نوع اول و دوم در تحلیل آماری نتایج
۴۴	۵-۲ تحلیل واریانس (ANOVA)
۴۸	۶-۲ آزمون مربع کای (χ^2)
۴۹	۷-۲ بررسی نرمال بودن توزیع داده‌ها و منحنی فراوانی تجمعی
۵۲	خلاصه فصل دوم
۵۴	خودآزمایی چهارگزینه‌ای فصل دوم
۵۶	خودآزمایی تشریحی فصل دوم
۵۹	فصل سوم: روش‌های درجه‌بندی
۵۹	هدف کلی
۵۹	هدف‌های یادگیری
۶۰	مقدمه
۶۱	۱-۳ درجه‌بندی
۶۲	۲-۳ نمودارهای درجه‌بندی
۶۴	۳-۳ ضریب همبستگی حاصل ضرب-گشتاور
۶۸	۴-۳ خط برگشت Y روی X و روش کمترین مربعات
۷۳	۵-۳ محاسبه غلظت نمونه مجهول
۷۶	۶-۳ روش‌های ارزیابی نیکویی برازش
۸۰	۷-۳ برازش منحنی
۸۴	۸-۳ ارقام شایستگی
۸۶	۹-۳ مقایسه دو روش با استفاده از خط برگشت
۹۱	خلاصه فصل سوم
۹۳	خودآزمایی چهارگزینه‌ای فصل سوم

خودآزمایی تشریحی فصل سوم.....	۹۵
فصل چهارم: مروری بر مفاهیم اولیهٔ جبر خطی.....	۹۹
هدف کلی.....	۹۹
هدف‌های یادگیری.....	۹۹
مقدمه.....	۱۰۰
۱-۴ استفاده از ماتریس داده‌ها به‌جای جدول داده‌ها.....	۱۰۰
۲-۴ بردارها.....	۱۰۱
۱-۲-۴ تعاریف.....	۱۰۱
۲-۲-۴ جمع و تفریق بردارها.....	۱۰۳
۳-۲-۴ ضرب اسکالر در بردار.....	۱۰۶
۴-۲-۴ ضرب بردار در بردار.....	۱۰۷
۵-۲-۴ طول و فاصله.....	۱۰۹
۳-۴ ماتریس‌ها.....	۱۱۳
۱-۳-۴ تعاریف.....	۱۱۳
۲-۳-۴ جمع و تفریق ماتریس‌ها.....	۱۱۶
۳-۳-۴ ضرب یک عدد اسکالر در ماتریس.....	۱۱۷
۴-۳-۴ ضرب ماتریس در ماتریس.....	۱۱۷
۵-۳-۴ معکوس یک ماتریس مربعی.....	۱۱۹
خلاصهٔ فصل چهارم.....	۱۲۹
خودآزمایی چهارگزینه‌ای فصل چهارم.....	۱۳۰
خودآزمایی تشریحی فصل چهارم.....	۱۳۲
فصل پنجم: برگشت چندمتغیره.....	۱۳۵
هدف کلی.....	۱۳۵
هدف‌های یادگیری.....	۱۳۵
مقدمه.....	۱۳۶
۱-۵ روش‌های هم مقیاس کردن داده‌ها.....	۱۳۷
۱-۱-۵ مرکزوار کردن داده‌ها.....	۱۳۷
۲-۱-۵ استاندارد کردن داده‌ها.....	۱۳۹
۲-۵ کوواریانس و همبستگی.....	۱۴۴
۳-۵ برگشت چندمتغیره.....	۱۵۱
۴-۵ کاربردهایی از برگشت خطی چندمتغیره در شیمی تجزیه.....	۱۵۴

۱۵۴.....	۵-۴-۱ مدل‌های ترکیب خطی.....
۱۵۸.....	۵-۴-۲ درجه‌بندی چندمتغیره.....
۱۶۶.....	۵-۴-۳ بهینه‌سازی چندمتغیره.....
۱۷۲.....	خلاصه فصل پنجم.....
۱۷۳.....	خودآزمایی چهارگزینه‌ای فصل پنجم.....
۱۷۵.....	خودآزمایی تشریحی فصل پنجم.....
۱۷۹.....	فصل ششم: کنترل کیفیت و نمودارهای آن.....
۱۷۹.....	هدف کلی.....
۱۷۹.....	هدف‌های آموزشی.....
۱۸۰.....	مقدمه.....
۱۸۳.....	۶-۱ نمودارهای کنترلی.....
۱۸۴.....	۶-۲ نمودار شیوه‌ها.....
۱۸۴.....	۶-۲-۱ نمودارهای شیوه‌ها برای مقادیر میانگین.....
۱۹۰.....	۶-۲-۲ نمودارهای شیوه‌ها برای گستره‌ها.....
۱۹۶.....	۶-۳-۱ متوسط طول اجرا و کاستی‌های نمودارهای شیوه‌ها.....
۱۹۸.....	۶-۳ نمودار کیوسام.....
۲۰۷.....	خلاصه فصل ششم.....
۲۰۹.....	خودآزمایی چهارگزینه‌ای فصل ششم.....
۲۱۱.....	خودآزمایی تشریحی فصل ششم.....
۲۱۳.....	پاسخ خودآزمایی‌های چهارگزینه‌ای.....
۲۱۵.....	پیوست الف- جدول‌های آماری.....
۲۱۹.....	کتاب‌نامه.....

پیشگفتار

رسالت اصلی دانش شیمی تجزیه اندازه‌گیری کمی انواع گونه‌ها (از ساده‌ترین عناصر تا پیچیده‌ترین ترکیبات زیستی) در طیف وسیعی از نمونه‌ها است. این اندازه‌گیری می‌تواند برای یک گونه منفرد انجام شود و یا به تعیین مقدار هم‌زمان چندین گونه در نمونه بیانجامد. در این راستا یک تجزیه‌گر در آزمایشگاه بایستی طی گام‌های پی‌درپی و مراحل گوناگون که هریک از آن‌ها ممکن است بارها و بارها تکرار شود و ساعت‌ها و یا حتی روزها به طول بیانجامد، عمل اندازه‌گیری را به انجام رساند و درنهایت حاصل تلاش‌های خود را به صورت یک مقدار کمی عددی گزارش دهد. به‌ویژه در آزمایشگاه‌های شیمی تجزیه امروزی با انواع دستگاه‌های مدرن و پیشرفته، تجزیه‌گر در هریک از مراحل اندازه‌گیری با انبوهی از داده‌های عددی و گرافیکی مواجه می‌شود که اگر توانایی تجزیه و تحلیل آن‌ها را نداشته باشد موجب سردرگمی او شده و درنهایت منجر به نتیجه‌گیری ناصحیح خواهد شد. استفاده از علم آمار در تلفیق با شیمی تجزیه به‌منظور شناسایی، طبقه‌بندی و تفسیر داده‌های حاصل می‌تواند به چنین سردرگمی‌هایی خاتمه بخشیده و منجر به ارائه نتیجه‌ای قابل‌اعتماد گردد. هدف از نگارش این کتاب آشنایی دانشجویان عزیز با اصول پایه و مقدماتی علم آمار و کاربرد آن در شیمی تجزیه است. این اصول و کاربرد آن‌ها به صورت مرحله‌به‌مرحله در قالب شش فصل در کتاب حاضر ارائه شده است.

در فصل اول ضمن معرفی متداول‌ترین مفاهیم و اصطلاحات آماری، به بحث در مورد انواع خطاها و انتشار آن‌ها در محاسبات پرداخته شده است و درنهایت چگونگی گزارش یک نتیجه قابل‌اعتماد در سطح اطمینان مشخص با استفاده از مفاهیم فاصله و حدود اطمینان ارائه شده است. در فصل دوم بر پایه استفاده از مفاهیم ارائه شده در فصل اول، به بررسی نوع خطای همراه با نتایج تجزیه‌ای پرداخته شده است. در این راستا انواع آزمون‌های معنی‌دار بودن و موارد کاربرد هریک از آن‌ها در تحلیل آماری نتایج مورد بحث قرار گرفته است. مفهوم معیارگیری تک متغیره و نحوه ترسیم نمودارهای آن در

فصل سوم ارائه شده است. به علاوه، کلیه مفاهیم قابل استخراج از این نمودارها شامل گستره خطی، حد تشخیص، حد کمی سازی و حساسیت در این فصل به طور کامل معرفی شده است. در فصل چهارم مروری بر مفاهیم اولیه جبر خطی جهت یادآوری اصول حاکم بر جبر ماتریس ها و بردارها ارائه شده است. در فصل پنجم ضمن آشنایی با داده های چندمتغیره، مفاهیم هم مقیاس کردن داده ها و معیار گیری چندمتغیره مطرح شده است. در نهایت، در فصل ششم چگونگی کنترل کیفیت فرایند با استفاده از نمودارهای کنترلی مورد بحث قرار گرفته است. علاوه بر این، در پایان کتاب پیوستی مربوط به جداول آماری آمده است.

در تدوین این کتاب کوشش نویسندگان بر این بوده است که تمامی مطالب با بیانی هرچه ساده تر و با ذکر مثال های متعدد مرتبط با رشته شیمی، به صورت قابل فهم و ملموس ارائه گردد. همچنین در پایان هر فصل تعدادی خودآزمایی چهارگزینه ای و خودآزمایی تشریحی مطرح شده است که دانشجویان عزیز می توانند جهت سنجش آموخته های خود از آن ها بهره گیرند. در پایان از کلیه دانشجویان، اساتید و صاحب نظران گرامی تقاضا می شود که با ارائه نظرات، پیشنهادات و ذکر ایرادهای احتمالی این کتاب ما را در ارتقای کیفیت آن در چاپ های بعدی یاری دهند. امید است مطالب این کتاب راهگشای مسیر علمی-پژوهشی دانشجویان و پژوهشگران گرامی باشد.

عزیزاله نژادعلی

Aziz_Nezhadali@pnu.ac.ir

ملیحه مجرب

Maliheh.Mojarrab@gmail.com

فصل اول

مفاهیم اولیه

هدف کلی

در این فصل دانشجو ضمن آشنایی با مفاهیم اولیه آمار، انواع تعاریف آن را فراگرفته و کاربرد آن‌ها را در آزمون‌های آماری می‌آموزد.

هدف‌های یادگیری

دانشجویان، پس از یادگیری مطالب این فصل باید بتوانند:

۱. تفاوت میان آمار توصیفی و آمار استنباطی را شرح دهید.
۲. انواع خطا را نام ببرید.
۳. انواع خطای نظام‌مند را توضیح دهید.
۴. تفاوت میان خطای تصادفی و خطای نظام‌مند را بیان کنید.
۵. انواع عواملی که جهت تعیین بزرگی خطاهای تصادفی به کار می‌روند را نام ببرید.
۶. تعریفی از مفاهیم دقت و صحت ارائه دهید و تفاوت میان این دو مفهوم را شرح دهید.
۷. تفاوت میان خطای مطلق و خطای نسبی را با ذکر روابط مربوطه بیان کنید.

۸. نحوه محاسبه میانگین و میانه را بیان کنید.
۹. مزیت استفاده از میانه نسبت به میانگین را شرح دهید.
۱۰. مفاهیم مُد، مُد اصلی و مُد فرعی را تعریف کنید.
۱۱. تعریفی از دامنه و انحراف از میانگین ارائه کنید.
۱۲. چگونگی محاسبه انحراف استاندارد، انحراف استاندارد نسبی، درصد انحراف استاندارد نسبی و واریانس را بیان کنید.
۱۳. نقش انحراف استاندارد را در درک صحیح دقت شرح دهید.
۱۴. قضیه حد مرکزی را شرح دهید.
۱۵. با ذکر روابط، مفهوم حدود اطمینان و فاصله اطمینان را توضیح دهید.
۱۶. ارقام بامعنی را تعریف کنید.
۱۷. مفهوم انتشار خطا در محاسبات را بیان کنید.
۱۸. روابط مربوط به محاسبه خطای انتشار یافته در انواع محاسبات را ذکر کنید.

مقدمه

آمار عموماً مترادفی برای ارقام و اطلاعات است. مثلاً چه تعداد کارمند در یک اداره یا چه تعداد دانش آموز در یک دبیرستان وجود دارد. آمار در واقع از اطلاعات و داده‌ها در یک گروه، سری و یا مجموعه، برای تجزیه و تفسیر آن‌ها و همچنین دستیابی به شرایط بهتر، استفاده می‌کند. به عبارت دیگر، از داده‌های موجود برای استخراج اطلاعات و دانش بیشتر استفاده می‌شود. آمار شاخه‌ای از ریاضیات کاربردی است که به دو بخش آمار توصیفی^۱ و آمار استنباطی^۲ تقسیم می‌شود. در آمار توصیفی کارها به صورت تنظیم و طبقه‌بندی داده‌ها به روش رسم (ترسیمی)، جدول و محاسبه مقادیر و کمیت‌هایی نظیر میانه، میانگین، مُد، واریانس، انحراف استاندارد، گستره، درصد خطا و غیره است. این طبقه‌بندی و اطلاعات صرفاً در مورد همان دسته و گروه کاربرد داشته و قابل تعمیم به گروه‌های مشابه نیست. در مقابل اگر به جای مطالعه کل اعضای جامعه (جمعیت) بخشی از آن (نمونه) با روش‌های صحیح نمونه‌گیری مطالعه شده و نتایج حاصل از آن به جامعه تعمیم داده شود، آمار استنباطی استفاده شده است. در یک تقسیم‌بندی دیگر آمار به دو

1. Descriptive statistic

2. Inferential statistic

بخش کلاسیک و بایاس^۱ تقسیم‌بندی می‌شود. در آمار کلاسیک ابتدا آزمایش انجام می‌شود و سپس فرض‌های آن آزمون می‌شود. در آمار بایاس ابتدا فرضی در نظر گرفته‌شده و سپس داده‌ها با آن منطبق می‌شوند. تفاوت‌های مشاهده‌شده بین فرض و داده‌های تجربی به خطاهای احتمالی در فرایند نمونه‌گیری، آزمایش و غیره برمی‌گردد.

۱-۱ انواع خطا در نتایج و داده‌ها

معمولاً در تحلیل‌های آماری، خطاهای موجود به سه بخش عمده، معین و نامعین تقسیم‌بندی می‌شوند. داده‌دارای خطاهای عمده یا درشت اغلب با تکرار یک یا دو بار آزمایش مشخص‌شده و حذف می‌شود. چنانچه در انتقال یک محلول شیمیایی از یک ظرف به ظرف دیگر بخشی از آن بیرون از ظرف ریخته شود، یک خطای عمده رخ داده است که با تکرار آزمایش کاملاً مشخص می‌شود که خطا بزرگ بوده است. به هنگام ثبت داده‌ها و یادداشت‌برداری اگر در یک عدد دو رقمی ممیز آن فراموش شود ده برابر خطا حاصل خواهد شد که از نوع خطاهای عمده است. فرض کنید در یادداشت یک نتیجه حجم‌سنجی به جای عدد ۲/۲ عدد ۲۲ ثبت شود، در این حالت ده برابر یا حدود ۱۰۰۰٪ خطا صورت گرفته است که با تکرار یک یا دو بار دیگر آزمایش کاملاً مشخص می‌شود که اشتباه عمده‌ای صورت گرفته است. مثال دیگر، سرریز یک آزمایش به هنگام افزودن ماده‌ای به ماده دیگر یا حرارت دادن یک محلول و حذف یک‌بار و ناخواسته بخش زیادی از محلول است. به دلیل تشخیص سریع و راحت این نوع خطا در بیشتر موارد شیمیدان‌ها از خطاهای عمده اسم نبرده و خطاها را به دو گروه معین (نظام‌مند) و نامعین (تصادفی) تقسیم‌بندی می‌کنند.

۱-۱-۱ خطاهای معین

خطاهای نظام‌مند (معین) که به آن‌ها خطای قابل محاسبه نیز گفته می‌شود در آزمایش مقداری ثابت داشته و جهت معینی دارند. جهت این خطاها (مثبت یا منفی) با تکرار تغییر نکرده و تابع خطای شخص، دستگاه و یا روش آزمایش است. این عوامل (شخص، دستگاه یا روش) می‌توانند انفرادی یا هم‌زمان باعث خطای معین شوند. در آزمایشگاه‌های شیمی ظروف حجمی مدرج بوده و عدد حجم بر روی آن نوشته می‌شود؛ اما ممکن است

این عدد با حجم واقعی آن ظرف، هم‌خوانی نداشته باشد (مثلاً حجم واقعی بیشتر از عدد نوشته‌شده بر روی آن باشد) و آزمایشگر هر چند بار با آن ظرف حجمی، محلول را تهیه کند حجم بیشتری از عدد نوشته‌شده روی ظرف برمی‌دارد و یک خطای معین دستگاهی خواهد داشت که همیشه جهت خطا مثبت است. به‌عنوان مثالی دیگر که خطای معین منفی داشته باشد، می‌توان از لامپ‌های دستگاه‌هایی که در اندازه‌گیری‌های کمی مورد استفاده قرار می‌گیرند نام برد. طول عمر لامپ‌ها (مثلاً لامپ یک دستگاه فرابنفش-مرئی) معمولاً تابع زمان روشن بودن و استفاده از آن‌هاست. چنانچه از یک لامپ زمان زیادی استفاده شود، آن لامپ ضعیف شده و شدت نشر حاصل از آن نسبت به مقدار اسمی آن کم می‌شود و چنانچه این لامپ درجه‌بندی نشود یک خطای معین منفی خواهد داشت. علاوه بر خطای معین دستگاهی ممکن است خطای نظام‌مند ناشی از خود آزمایشگر باشد. برخی از انسان‌ها ممکن است ذاتاً نارسایی بینایی داشته باشند و به‌هنگام مشاهده تغییر رنگ یک محلول در نقطه ختم عمل تشخیص زودهنگام یا دیرهنگام رنگ را داشته باشند. بدیهی است در این نوع خطا نیز، به‌عنوان نمونه‌ای از خطای معین فردی، تکرار آزمایش تغییری در نتیجه ایجاد نخواهد کرد. در نوع سوم خطاهای معین (خطای روش) عوامل مختلف دستور کار آزمایش نظیر ناپایداری واکنش‌گرها، انحلال ذاتی نمک حاصل از آزمایش و موارد مشابه می‌تواند باعث خطای مثبت یا منفی معین در نتیجه آزمایش شود. خطاهای معین می‌توانند به‌صورت متناسب یا ثابت باشند. در نوع ثابت، میزان خطا تابع مقدار کمیت مورد اندازه‌گیری نیست، مثلاً از بین رفتن مقداری از رسوب کم‌محلول در اثر حلالیت، تابع کم یا زیاد شدن مقدار رسوب نبوده و فقط حاصل ضرب انحلال آن نمک کم‌محلول (K_{sp}) در این مورد تأثیرگذار است. در مقابل خطای متناسب بستگی به مقدار کمیت مورد اندازه‌گیری دارد. برای مثال، اگر محلول‌های لازم برای تجزیه توسط حل کردن نمونه جامد در آب تهیه شود و این نمونه یک مزاحم به همراه داشته باشد، افزایش غلظت ماده مورد تجزیه (آنالیت) با افزایش جرم برداشته‌شده (جامد) برای حجم ثابتی از محلول، باعث افزایش غلظت ماده مزاحم شده و افزایش خطای معین را در پی دارد.

۱-۲ خطاهای نامعین

خطاهای نامعین (تصادفی) منشأ خاصی ندارند و احتمال خطای نامعین مثبت و منفی یکسان است. هرچه اندازه این خطا کوچک‌تر باشد احتمال بروز آن بیشتر است. خطاهای تصادفی از کنترل آزمایشگر خارج است. در اینجا نیز مشابه خطاهای معین از هر سه مورد فرد، دستگاه و روش انتظار خطا می‌رود. عواملی نظیر تغییر دما، فشار و یا هوای آزمایشگاه می‌تواند باعث بروز خطاهای تصادفی شود. چون خطاهای نامعین جهت مشخصی ندارند تکرار آزمایش‌ها باعث کاهش خطاهای برخی از عوامل می‌شود؛ بنابراین، با تکرار بیشتر آزمایش‌ها و معدل‌گیری از جواب، نتایج بهتری حاصل می‌شود. در عمل با کمک عواملی نظیر دقت^۱، انحراف استاندارد^۲ و واریانس^۳ می‌توان به بزرگی خطای تصادفی پی برد.

۱-۲ برخی اصطلاحات آماری

۱-۲-۱ دقت و صحت^۴ داده‌ها

گرچه در بحث‌های روزانه و معمول دقت و صحت به‌طور مترادف استفاده می‌شوند، ولی در آمار، به‌طور متمایز از یکدیگر تعریف می‌شوند. صحت عبارت از نزدیکی یک مقدار اندازه‌گیری شده یا محاسبه‌شده به مقدار پذیرفته‌شده (واقعی) است، مثلاً اگر وزن واقعی یک کمپلکس سنتز شده ۲/۸۸ گرم باشد و دانشجوی ۱ و ۲ در آزمایشگاه به ترتیب ۲/۸۲ و ۲/۹۱ گرم را برای سنتزهای خود ثبت کرده باشند، جواب دانشجوی ۲ صحت بیشتری نسبت به دانشجوی ۱ دارد (چون به وزن واقعی نزدیک‌تر است). دقت، عبارت از نزدیکی اندازه‌گیری‌های تکراری از یک کمیت مشابه با یکدیگر است. اگر دانشجویان ۱ و ۲ هرکدام ۵ بار آزمایش خود را تکرار کنند و نتایج آن‌ها به ترتیب برای اولی ۲/۸۲، ۲/۸۹، ۲/۸۸، ۲/۹۰ و ۲/۸۸ باشد و برای دومی ۲/۹۱، ۲/۸۱، ۲/۶۲، ۲/۹۵ و ۲/۹۷ باشد، دقت نتایج ثبت‌شده توسط اولی بیشتر از دومی است. در اینجا پراکنش داده‌های اولی کمتر از

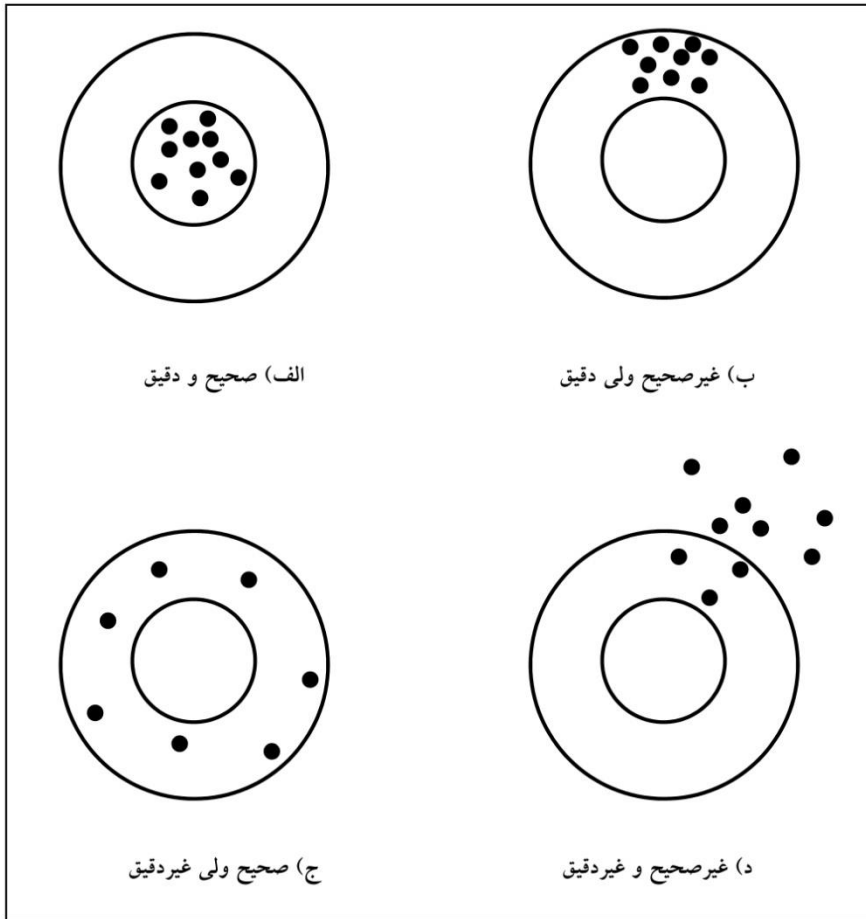
1. Precision

2. Standard deviation (SD)

3. Variance

4. Accuracy

دومی بوده و به یکدیگر نزدیک تر است. برای تمایز بهتر بین صحت و دقت، وضعیت تیراندازی به یک سیبل در شکل ۱-۱، ارائه شده است.



شکل ۱-۱. مقایسه صحت و دقت در یک تیراندازی.

الف) صحیح و دقیق ب) غیر صحیح ولی دقیق، ج) صحیح ولی غیردقیق و د) غیر صحیح و غیردقیق.

همانگونه که شکل نشان می دهد در قسمت الف تیرهای رهاشده هم به هدف خورده و هم این که کنار یکدیگر هستند؛ بنابراین، فرد الف فردی دقیق و صحیح است. در ب علی رغم این که تیرها کنار هم هستند ولی از هدف دور شده اند که می توان گفت فرد ب فردی دقیق ولی غیر صحیح است. فرد ج تیراندازی نسبتاً صحیح داشته ولی دقیق نیست و تیراندازی فرد د نه صحیح بوده و نه دقیق است.

۲-۲-۱ خطای مطلق^۱

تفاوت میان مقدار اندازه‌گیری شده (x_i) و مقدار واقعی (x_t)، خطای مطلق آن اندازه‌گیری نامیده می‌شود (رابطه ۱-۱). چنانچه مقدار خطای مطلق مثبت باشد، مقدار اندازه‌گیری شده بیشتر از مقدار واقعی است و اگر منفی باشد، مقدار اندازه‌گیری شده کمتر از مقدار واقعی است.

$$E = x_i - x_t \quad (1-1)$$

۳-۲-۱ خطای نسبی

مقدار خطا و درصد آن به صورت نسبی و درصدی هم بیان می‌شود. از تقسیم مقدار خطای مطلق به مقدار واقعی، خطای نسبی و متعاقب آن درصد خطای نسبی به دست می‌آید که به صورت رابطه‌های (۲-۱) و (۳-۱)، بیان می‌شود.

$$E_r = \frac{x_i - x_t}{x_t} \quad (2-1)$$

$$E_r\% = \frac{x_i - x_t}{x_t} \times 100 \quad (3-1)$$

که E_r و $E_r\%$ به ترتیب خطای نسبی و درصد خطای نسبی هستند.

❖ مثال ۱-۱

مقدار یون نقره در یک سنگ معدن ۶/۷ ppm اندازه‌گیری شده است. اگر مقدار واقعی یون نقره در این نمونه ۶/۹ ppm باشد، خطای مطلق، خطای نسبی و درصد خطای نسبی اندازه‌گیری را محاسبه کنید.

$$E = 6/7 - 6/9 = -0/2 \text{ ppm}$$

پاسخ:

$$E_r = \frac{6/7 - 6/9}{6/9} = -0/029$$

$$E_r\% = \frac{6/7 - 6/9}{6/9} \times 100 = -2/9\%$$

تمرین ۱-۱

توسط دستگاه جذب اتمی، مقدار نیکل در آلیاژی g ۰/۱۱ اندازه‌گیری می‌شود. اگر مقدار واقعی نیکل در نمونه آلیاژ g ۰/۱۲ باشد، درصد خطای نسبی را محاسبه کنید.
پاسخ: ۸/۳٪

۱-۲-۴ میانگین حسابی^۱

مهم‌ترین معیار متمایل به مرکز داده‌ها، میانگین حسابی یا همان میانگین است که معمولاً به آن معدل^۲ گفته می‌شود. هر اندازه‌گیری از یک جمعیت به صورت یک ارزش x_i معرفی می‌شود (مشاهده مربوط به فرد i ام برای متغیر x). اندیس i می‌تواند هر ارزش صحیح تا N باشد که N تعداد کل ارزش‌های x در جمعیت یا اندازه جمعیت است. میانگین جمعیت با μ نمایش داده شده و به صورت زیر تعریف می‌شود.

$$\mu = \frac{\sum_{i=1}^N x_i}{N} \quad (4-1)$$

مناسب‌ترین تخمین از میانگین جمعیت (μ)، میانگین نمونه است که به صورت $\bar{x}$ بیان می‌شود (رابطه ۵-۱).

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (5-1)$$

❖ مثال ۱-۲

تعداد ۲۴ نمونه سنگ معدن طلا تجزیه شده و نتایج زیر برحسب ppm به دست آمده است. میانگین غلظت طلا در این سنگ معدن چقدر است؟

۳/۸، ۳/۹، ۳/۹، ۴/۰، ۳/۹، ۴/۰، ۴/۰، ۴/۰، ۴/۱، ۴/۱، ۴/۱، ۴/۲، ۴/۳، ۴/۳، ۴/۴، ۴/۵، ۴/۵، ۳/۸، ۳/۷، ۳/۶، ۳/۶، ۳/۵، ۳/۳ و ۴/۲.

پاسخ:

با کمک معادله (۱-۵) مقدار میانگین به صورت زیر محاسبه می شود:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{۹۵}{۲۴} = ۳/۹۶ \text{ ppm}$$

۱-۲-۵ میانه

در یک مجموعه داده مرتب، میانه مقادیر اندازه گیری شده در وسط داده ها قرار دارد. برای تعیین میانه پس از مرتب کردن داده ها، اگر تعداد داده ها فرد باشد عدد وسط انتخاب می شود و اگر تعداد آن ها زوج باشد میانگین دو عدد وسط مقدار میانه است. میانه اطلاعات کمتری نسبت به میانگین ارائه می دهد؛ زیرا برای محاسبه آن از ارزش واقعی هر مشاهده استفاده نمی شود و فقط رتبه هر مشاهده در نظر گرفته می شود. در برخی حالت ها میانه نسبت به میانگین دارای مزیت است. اگر تعداد داده ها بی نهایت بزرگ یا بی نهایت کوچک باشد، روی میانه اثر ندارد (فقط رتبه آن ها مهم است) ولی به میزان زیادی میانگین را تحت تأثیر قرار می دهند. در چنین حالتی میانه نمونه، یک آماره مقاوم^۱ نامیده می شود.

۱-۲-۶ مَد

در یک مجموعه از داده ها، داده ای که بیشتر از همه تکرار شود، مَد نامیده می شود. به عبارت دیگر، مَد مقادیری از نتایج است که تراکم نسبی بیشتری دارند. توزیعی که در آن نتایج مختلف فراوانی یکسانی دارند، فاقد مَد است. اگر در توزیعی دو نتیجه دارای تکرارهای نزدیک به هم باشد، بنا به فراوانی تکرارها، آن دو نتیجه مدهای اصلی و فرعی نامیده می شوند.

۸-۲-۱ انحراف از میانگین^۱

میانگین معیار مفید متمایل به مرکز مجموعه داده‌ها است. پراکندگی داده‌ها می‌تواند به صورت انحراف از میانگین بیان شود. مجموع کلیه انحراف‌ها از میانگین که به صورت $\sum (x_i - \bar{x})$ بیان می‌شود همیشه مساوی صفر است. بدین منظور برای متوسط انحراف از میانگین نمونه از قدر مطلق استفاده شده و می‌توان رابطه (۷-۱) را نوشت.

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n} = \text{متوسط انحراف از میانگین نمونه} \quad (۷-۱)$$

۹-۲-۱ انحراف استاندارد

در بیشتر کارهای کمی مخصوصاً در شیمی تجزیه، برای بیان گستردگی نتایج در اطراف میانگین از انحراف استاندارد استفاده می‌شود. انحراف استاندارد با s نمایش داده شده و برای n بار اندازه‌گیری تکراری، می‌تواند به صورت یکی از روابط (۸-۱) تا (۱۰-۱) تعریف شود.

$$s = \sqrt{\frac{\left(\sum_{i=1}^n x_i^2 \right) - n\bar{x}^2}{n-1}} \quad (۸-۱)$$

$$s = \sqrt{\frac{\sum x_i^2 - \left[\left(\sum x_i \right)^2 / n \right]}{n-1}} \quad (۹-۱)$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \quad (۱۰-۱)$$

این کمیت در برخی موارد به صورت SD نیز ارائه می‌شود و در برخی کتب آماری جذر میانگین مربع انحراف^۲ یا جذر میانگین مربع^۳ ذکر می‌شود. از تقسیم انحراف

1. Mean deviation
2. Root mean square deviation
3. Root mean square