بسمه تعالی

موضوع : مبانی آمار و احتمال برای ورود به علم داده و تحلیل داده

بخش اخر:

- چرا نمونه‌گیری می‌کنیم؟
- خطای نمونه‌گیری
- برآورد
- مفاهیم
- پارامتر
- آماره
- توزیع نمونه‌گیری
- فاصله اطمینان
- مفهوم بدون ورود به اثبات
- آزمون فرض
- فرض صفر (H0)
- فرض مقابل (H1)
- p-value به زبان ساده
- آزمون z
- آزمون t

در این پارت از اموزش به بررسی **مقدماتی آمار استنتاجی** می پردازیم .

در اموزش های قبلی راجب **به طور کامل درباره آمار توصیفی** صحبت کردیم.

تا اینجا بیشتر درباره آمار توصیفی صحبت کردیم؛ یعنی داده‌هایی که در اختیار داریم را جمع‌آوری، خلاصه و توصیف می‌کنیم. اما سؤال مهمی که در بسیاری از مسائل آماری مطرح می‌شود این است که اگر فقط بخشی از داده‌ها را در اختیار داشته باشیم، آیا می‌توانیم درباره کل جامعه هم نتیجه‌گیری کنیم؟

اینجاست که وارد مبحث آمار استنباطی یا Inferential Statistics می‌شویم. در آمار استنباطی، ما معمولاً از یک نمونه استفاده می‌کنیم و بر اساس اطلاعات نمونه، درباره جامعه آماری نتیجه‌گیری می‌کنیم.

چرا نمونه‌گیری می‌کنیم؟

فرض کنید یک شرکت ۱۰۰ هزار مشتری دارد و می‌خواهیم بدانیم میانگین میزان رضایت مشتریان چقدر است. از نظر تئوری می‌توانیم از همه ۱۰۰ هزار نفر اطلاعات بگیریم، اما در عمل این کار ممکن است بسیار زمان‌بر، پرهزینه یا حتی غیرممکن باشد.

بنابراین، مثلاً ۵۰۰ نفر از مشتریان را انتخاب می‌کنیم و میزان رضایت آن‌ها را بررسی می‌کنیم. این ۵۰۰ نفر نمونه هستند و تمام ۱۰۰ هزار مشتری، جامعه ما محسوب می‌شوند. هدف این است که نمونه‌ای انتخاب کنیم که تا حد امکان نماینده مناسبی از جامعه باشد.

خطای نمونه‌گیری

اما یک نکته مهم وجود دارد. حتی اگر نمونه را به‌درستی انتخاب کنیم، میانگین نمونه لزوماً دقیقاً برابر با میانگین واقعی جامعه نیست. مثلاً فرض کنید میانگین واقعی رضایت مشتریان ۷۵ باشد، اما در نمونه‌ای که انتخاب کرده‌ایم میانگین ۷۳ به دست بیاید. این اختلاف، بخشی از خطای نمونه‌گیری یا Sampling Error است. بنابراین باید توجه داشته باشیم که وقتی از نمونه برای نتیجه‌گیری درباره جامعه استفاده می‌کنیم، همیشه مقداری عدم‌قطعیت وجود دارد. آمار استنباطی در واقع ابزارهایی در اختیار ما قرار می‌دهد تا این عدم‌قطعیت را اندازه‌گیری و مدیریت کنیم.

برآورد

یکی از کاربردهای اصلی آمار استنباطی، برآورد پارامترهای جامعه است.برای مثال، اگر میانگین واقعی درآمد تمام کارکنان یک شرکت را ندانیم، می‌توانیم از میانگین درآمد یک نمونه از کارکنان برای تخمین آن استفاده کنیم. در اینجا دو نوع برآورد را می‌توان مطرح کرد:

برآورد نقطه‌ای و برآورد فاصله‌ای.

در برآورد نقطه‌ای، یک مقدار را به‌عنوان تخمین ارائه می‌کنیم؛ مثلاً می‌گوییم میانگین درآمد جامعه حدود ۳۰ میلیون تومان است.

اما در برآورد فاصله‌ای، به‌جای یک عدد، یک بازه ارائه می‌کنیم؛ مثلاً می‌گوییم میانگین درآمد جامعه احتمالاً در بازه ۲۸ تا ۳۲ میلیون تومان قرار دارد.

این مفهوم ما را به فاصله اطمینان می‌رساند.

پارامتر و آماره

قبل از ادامه، باید تفاوت پارامتر و آماره را بدانیم.پارامتر یک ویژگی عددی مربوط به کل جامعه است. مثلاً میانگین واقعی درآمد تمام کارکنان یک شرکت، یک پارامتر است و معمولاً با نماد μ نشان داده می‌شود. اما اگر از جامعه یک نمونه انتخاب کنیم و میانگین درآمد همین نمونه را محاسبه کنیم، این مقدار یک آماره یا Statistic است و معمولاً با $\bar{x}$ نشان داده می‌شود.

بنابراین یک قانون ساده را به خاطر بسپارید:

پارامتر ⟵ مربوط به جامعه

آماره ⟵ مربوط به نمونه

ما معمولاً پارامتر جامعه را نمی‌دانیم و از آماره نمونه برای برآورد آن استفاده می‌کنیم.

توزیع نمونه‌گیری

حالا یک مفهوم مهم دیگر داریم که به آن توزیع نمونه‌گیری یا Sampling Distribution می‌گوییم. فرض کنید از یک جامعه، تعداد زیادی نمونه با حجم یکسان انتخاب کنیم و میانگین هر نمونه را به دست بیاوریم. میانگین نمونه اول ممکن است ۷۲ باشد، نمونه دوم ۷۵، نمونه سوم ۷۳ و نمونه چهارم ۷۶. بنابراین میانگین نمونه‌ها با یکدیگر تفاوت دارند.

اگر این میانگین‌ها را در کنار هم بررسی کنیم، یک توزیع از آماره موردنظر خواهیم داشت که به آن توزیع نمونه‌گیری می‌گوییم. این مفهوم اهمیت زیادی دارد، چون پایه بسیاری از روش‌های فاصله اطمینان و آزمون فرض است.

فاصله اطمینان

یکی از مهم‌ترین مفاهیم آمار استنباطی، Confidence Intervalیا فاصله اطمینان است. فرض کنید میانگین نمونه ما ۷۵ به دست آمده است.

آیا می‌توانیم بگوییم میانگین واقعی جامعه دقیقاً ۷۵ است؟ خیر.

چون این مقدار از یک نمونه به دست آمده و ممکن است با مقدار واقعی جامعه تفاوت داشته باشد. بنابراین به‌جای ارائه یک مقدار قطعی، می‌توانیم یک بازه ارائه کنیم. مثلاً:

میانگین نمونه = ۷۵

فاصله اطمینان ۹۵ درصد = ۷۲ تا ۷۸

در این حالت، به‌جای اینکه فقط بگوییم میانگین جامعه ۷۵ است، یک بازه برای پارامتر جامعه ارائه می‌کنیم.

سطح اطمینان چیست؟

یکی از سطوح رایج، ۹۵ درصد است. سطح اطمینان نشان‌دهنده میزان اطمینانی است که روش برآورد ما در بلندمدت دارد. برای مثال، اگر یک روش را برای نمونه‌گیری‌های متعدد تکرار کنیم و برای هر نمونه یک فاصله اطمینان ۹۵ درصدی بسازیم، در شرایط ایده‌آل حدود ۹۵ درصد از این فاصله‌ها پارامتر واقعی جامعه را دربر می‌گیرند.

یک نکته مهم:

فاصله اطمینان ۹۵ درصد را نباید به این شکل تفسیر کنیم که «احتمال اینکه پارامتر واقعی داخل این بازه خاص باشد، ۹۵ درصد است.« در چارچوب کلاسیک، پارامتر جامعه یک مقدار ثابت است و این روش ساخت فاصله است که دارای سطح پوشش ۹۵ درصدی است. برای دوره مقدماتی، فعلاً همین مفهوم برای ما کافی است و وارد اثبات ریاضی آن نمی‌شویم.

آزمون فرض

گاهی هدف ما تخمین یک مقدار نیست، بلکه می‌خواهیم یک ادعا یا فرض آماری را بررسی کنیم. برای مثال، یک کارخانه ادعا می‌کند میانگین عمر یک محصول ۵ سال است. ما تعدادی محصول را به‌عنوان نمونه انتخاب می‌کنیم و عمر آن‌ها را اندازه‌گیری می‌کنیم. حالا می‌خواهیم ببینیم آیا اطلاعات نمونه با این ادعا سازگار است یا شواهد آماری کافی برای رد آن داریم. به این فرآیند آزمون فرض یا Hypothesis Testing می‌گوییم.

فرض صفر و فرض مقابل

در آزمون فرض معمولاً دو فرض داریم:

فرض صفر یاH_0 فرض صفر معمولاً بیان می‌کند که تفاوت یا اثر موردنظر وجود ندارد یا مقدار پارامتر با مقدار مشخصی برابر است. مثلاً:

$H_0: \mu = ۵$

یعنی میانگین عمر جامعه برابر با ۵ سال است.

فرض مقابل یاH_1

فرض مقابل بیان می‌کند که ادعای فرض صفر برقرار نیست.

مثلاً:

$H_1: \mu \neq ۵$

یعنی میانگین عمر جامعه با ۵ سال تفاوت دارد. البته فرض مقابل می‌تواند یک‌طرفه هم باشد؛ مثلاً:

$H_1: \mu > ۵$

یا:

$H_1: \mu < ۵$

انتخاب نوع فرض مقابل باید بر اساس سؤال پژوهش و قبل از مشاهده نتیجه آزمون انجام شود.

سطح معناداری

حالا به یکی از مفاهیم مهم می‌رسیم: سطح معناداری یا Significance Level که معمولاً با حرف آلفا یا α نشان داده می‌شود. یکی از مقادیر رایج برای α برابر با ۰/۰۵ است. اگر α را برابر ۰/۰۵ در نظر بگیریم، در واقع قبل از انجام آزمون مشخص کرده‌ایم که چه مقدار ریسک برای رد اشتباه فرض صفر را می‌پذیریم. در آزمون فرض ممکن است تصمیم آماری ما با واقعیت متفاوت باشد.

دو حالت مهم وجود دارد:

خطای نوع اول یاType I Error

خطای نوع اول زمانی اتفاق می‌افتد که فرض صفر در واقع درست باشد، اما ما آن را رد کنیم.

به زبان ساده:

فرض صفر درست است، ولی ما اشتباهاً آن را رد کرده‌ایم.

احتمال این خطا با α ارتباط دارد.

اگر سطح معناداری را ۰/۰۵ انتخاب کنیم، سطح خطای نوع اول تعیین‌شده در آزمون ۵ درصد است.

خطای نوع دوم یاType II Error

خطای نوع دوم زمانی اتفاق می‌افتد که فرض صفر در واقع نادرست باشد، اما ما نتوانیم آن را رد کنیم.

یعنی:

فرض صفر غلط است، ولی شواهد کافی برای رد آن پیدا نکرده‌ایم.

احتمال خطای نوع دوم را معمولاً با β نشان می‌دهیم.

در مقابل، مقدار $۱ - \beta$ را توان آزمون یا Power می‌نامیم.

توان آزمون نشان می‌دهد آزمون ما با چه احتمالی می‌تواند یک اثر یا تفاوت واقعی را شناسایی کند.

p-valueچیست؟

حالا به یکی از پرکاربردترین مفاهیم آزمون فرض یعنی p-value می‌رسیم.

فرض کنید فرض صفر را درست در نظر بگیریم.

p-valueبه زبان ساده نشان می‌دهد که اگر فرض صفر درست باشد، مشاهده نتیجه‌ای به اندازه نتیجه مشاهده‌شده یا شدیدتر، چقدر با این فرض سازگار است.

بنابراین:

p-valueکوچک ← داده‌های مشاهده‌شده با H_0 سازگاری کمتری دارند.

p-valueبزرگ ← داده‌های مشاهده‌شده شواهد کمتری برای رد H_0 فراهم می‌کنند.

یک نکته بسیار مهم این است که p-value احتمال درست بودن فرض صفر نیست.

یعنی اگر p-value برابر با ۰/۰۳ باشد، نمی‌توانیم بگوییم احتمال درست بودن H_0 برابر با ۳ درصد است.

این یکی از اشتباهات رایج در تفسیر p-value است.

معناداری آماری

حالا مفهوم معناداری آماری یا Statistical Significance را می‌توانیم توضیح دهیم.

قبل از آزمون، معمولاً یک سطح معناداری، مثلاً ۰/۰۵، تعیین می‌کنیم.

سپس p-value را با α مقایسه می‌کنیم.

اگر:

$p\text{-value} < 0.05$

باشد، نتیجه را از نظر آماری معنادار در نظر می‌گیریم و فرض صفر را رد می‌کنیم.

اگر:

$p\text{-value} \geq 0.05$

باشد، شواهد آماری کافی برای رد فرض صفر نداریم.

البته باید دقت کنیم که «معناداری آماری» الزاماً به معنی اهمیت عملی یا اقتصادی نیست.

مثلاً ممکن است یک تفاوت بسیار کوچک از نظر آماری معنادار شود، اما از نظر عملی اهمیت چندانی نداشته باشد.

بنابراین در تحلیل آماری بهتر است علاوه برp-value ، به اندازه اثر، فاصله اطمینان و اهمیت عملی نتیجه نیز توجه کنیم.

آزمونZ

یکی از آزمون‌های معروف، آزمون Z است.

آزمون Z در شرایط مشخصی برای بررسی فرضیه درباره میانگین یا نسبت استفاده می‌شود.

در حالت کلاسیک، وقتی انحراف معیار جامعه مشخص باشد، آزمون Z برای میانگین کاربرد دارد.

برای مثال، فرض کنید بخواهیم بررسی کنیم آیا میانگین وزن محصولات یک کارخانه برابر با مقدار مشخصی است یا خیر و انحراف معیار جامعه را می‌دانیم.

در این شرایط می‌توانیم از آزمون Z استفاده کنیم.

در آزمونZ ، فاصله مقدار مشاهده‌شده از مقدار فرض‌شده را نسبت به خطای استاندارد می‌سنجیم و در نهایت می‌توانیم آماره آزمون و p-value را به دست بیاوریم.

آزمونt

آزمون t نیز برای آزمون درباره میانگین بسیار پرکاربرد است.

تفاوت مهم این است که وقتی انحراف معیار جامعه را نمی‌دانیم، معمولاً از انحراف معیار نمونه استفاده می‌کنیم و در این شرایط آزمون t کاربرد پیدا می‌کند.

آزمون t به‌خصوص در نمونه‌های کوچک اهمیت زیادی دارد.

آزمون t خودش انواع مختلفی دارد؛ مثلاً:

One-Sample t-testبرای مقایسه میانگین یک نمونه با یک مقدار مشخص

Independent Samples t-testبرای مقایسه میانگین دو گروه مستقل

Paired Samples t-testبرای مقایسه دو اندازه‌گیری وابسته یا جفت‌شده

در ادامه مباحث آماری، بسته به مسئله می‌توان از هرکدام از این آزمون‌ها استفاده کرد.

یک مثال جامع برای درک آمار استنباطی

حالا که با مفاهیم اصلی آمار استنباطی آشنا شدیم، بیایید همه این مفاهیم را در قالب یک مثال واقعی کنار هم قرار دهیم.

فرض کنید یک شرکت تولیدکننده لامپ ادعا می‌کند:

»میانگین عمر لامپ‌های تولیدی این شرکت ۱۰۰۰ ساعت است«.

ما می‌خواهیم بررسی کنیم که آیا داده‌های واقعی با این ادعا سازگار هستند یا خیر.

مرحله اول: جامعه و نمونه

در این مسئله، تمام لامپ‌هایی که شرکت تولید می‌کند، جامعه آماری ما هستند.

فرض کنید شرکت در طول یک سال یک میلیون لامپ تولید کرده است.

واضح است که ما نمی‌توانیم عمر هر یک میلیون لامپ را اندازه‌گیری کنیم؛ چون برای اندازه‌گیری عمر لامپ باید آن را تا زمان سوختن روشن نگه داریم و این کار بسیار زمان‌بر است.

بنابراین از بین این جامعه بزرگ، مثلاً ۳۰ لامپ را به‌صورت تصادفی انتخاب می‌کنیم.

این ۳۰ لامپ، نمونه آماری ما هستند.

حالا عمر این ۳۰ لامپ را اندازه‌گیری می‌کنیم.

فرض کنید میانگین عمر این ۳۰ لامپ برابر با:

۹۷۰ ساعت

به دست آمده است.

در اینجا یک سؤال مهم داریم:

آیا چون میانگین نمونه ۹۷۰ ساعت شده، پس میانگین واقعی تمام لامپ‌های شرکت هم ۹۷۰ ساعت است؟

خیر.

این مقدار فقط یک آماره نمونه است و ما می‌خواهیم با استفاده از آن درباره پارامتر جامعه نتیجه‌گیری کنیم.

مرحله دوم: آماره و پارامتر

در این مثال:

میانگین واقعی عمر تمام لامپ‌های شرکت، یک پارامتر جامعه است و آن را با μ نشان می‌دهیم.

اما میانگین ۹۷۰ ساعتی که از ۳۰ لامپ به دست آورده‌ایم، یک آماره نمونه است و با $\bar{x}$ نشان داده می‌شود.

پس:

→ μ میانگین جامعه

→ $\bar{x}$ میانگین نمونه

از آنجا که μ را نمی‌دانیم، از $\bar{x}$ برای برآورد آن استفاده می‌کنیم.

مرحله سوم: خطای نمونه‌گیری

فرض کنید میانگین واقعی عمر تمام لامپ‌ها در واقع ۹۸۰ ساعت باشد.

اما میانگین نمونه ما ۹۷۰ ساعت شده است.

پس بین میانگین نمونه و میانگین واقعی جامعه اختلاف وجود دارد.

این اختلاف ناشی از نمونه‌گیری، بخشی از خطای نمونه‌گیری است.

اگر نمونه دیگری از ۳۰ لامپ انتخاب کنیم، ممکن است میانگین نمونه ۹۷۵ ساعت شود و اگر نمونه دیگری انتخاب کنیم، شاید میانگین ۹۶۵ ساعت به دست بیاید.

بنابراین آماره نمونه از یک نمونه به نمونه دیگر تغییر می‌کند.

اینجاست که مفهوم توزیع نمونه‌گیری اهمیت پیدا می‌کند.

مرحله چهارم: فاصله اطمینان

حالا فرض کنید از اطلاعات نمونه، فاصله اطمینان ۹۵ درصدی را محاسبه کرده‌ایم و به نتیجه زیر رسیده‌ایم:

$۹۴۰ \leq \mu \leq ۱۰۰۰$

یعنی فاصله اطمینان ۹۵ درصدی برای میانگین عمر لامپ‌ها برابر با ۹۴۰ تا ۱۰۰۰ ساعت است.

حالا چه چیزی از این نتیجه متوجه می‌شویم؟

ما دیگر فقط نمی‌گوییم میانگین عمر لامپ‌ها ۹۷۰ ساعت است.

بلکه یک بازه برای میانگین جامعه ارائه می‌کنیم.

نکته مهم این است که این بازه عدم‌قطعیت ناشی از نمونه‌گیری را هم در نظر می‌گیرد.

از طرف دیگر، می‌بینیم که مقدار ادعایی شرکت، یعنی ۱۰۰۰ ساعت، در انتهای این فاصله قرار دارد.

اما برای اینکه دقیق‌تر درباره ادعای شرکت تصمیم بگیریم، می‌توانیم از آزمون فرض استفاده کنیم.

مرحله پنجم: تعریف فرض صفر و فرض مقابل

ادعای شرکت این است که میانگین عمر لامپ‌ها ۱۰۰۰ ساعت است.

پس می‌توانیم فرض صفر را به شکل زیر تعریف کنیم:

$H_0: \mu = ۱۰۰۰$

یعنی میانگین واقعی عمر لامپ‌ها برابر با ۱۰۰۰ ساعت است.

حالا فرض مقابل را تعریف می‌کنیم:

$H_1: \mu \neq ۱۰۰۰$

یعنی میانگین واقعی عمر لامپ‌ها با ۱۰۰۰ ساعت تفاوت دارد.

در اینجا چون می‌خواهیم بررسی کنیم آیا میانگین با ۱۰۰۰ تفاوت دارد یا خیر، آزمون ما دوطرفه است.

یعنی هم کمتر بودن میانگین از ۱۰۰۰ و هم بیشتر بودن آن از ۱۰۰۰ برای ما اهمیت دارد.

مرحله ششم: سطح معناداری

قبل از انجام آزمون، باید مشخص کنیم چه سطحی از خطا را می‌پذیریم.

فرض کنید سطح معناداری را برابر با:

$\alpha = 0.05$

در نظر بگیریم.

این یعنی ما سطح خطای نوع اول را ۵ درصد در نظر گرفته‌ایم.

حالا خطای نوع اول دقیقاً چیست؟

مرحله هفتم: خطای نوع اول

فرض کنید در واقع میانگین عمر تمام لامپ‌ها واقعاً ۱۰۰۰ ساعت باشد.

یعنی H_0 در واقع درست باشد.

اما ما بر اساس داده‌های نمونه، H_0را رد کنیم و بگوییم:

»شواهد آماری نشان می‌دهد میانگین عمر لامپ‌ها با ۱۰۰۰ ساعت تفاوت دارد«.

اگر این تصمیم اشتباه باشد، دچار خطای نوع اول شده‌ایم.

یعنی:

H_0درست است، اما ما آن را رد کرده‌ایم.

سطح معناداری ۰/۰۵، آستانه‌ای است که برای کنترل احتمال این نوع خطا در روش آزمون تعیین می‌کنیم.

مرحله هشتم: خطای نوع دوم

حالا حالت برعکس را در نظر بگیریم.

فرض کنید میانگین واقعی عمر لامپ‌ها در حقیقت ۹۰۰ ساعت باشد.

بنابراینH_0 ، یعنی $\mu = $ »۱۰۰۰«، درست نیست.

اما نمونه‌ای که انتخاب کرده‌ایم به اندازه کافی قوی نیست که این تفاوت را نشان دهد و ما H_0 را رد نمی‌کنیم.

در این حالت، خطای نوع دوم اتفاق افتاده است.

یعنی:

H_0در واقع نادرست است، اما ما نتوانسته‌ایم آن را رد کنیم.

پس خیلی خلاصه:

خطای نوع اول:

H_0درست است ⟵ ولی ما آن را رد می‌کنیم.

خطای نوع دوم:

H_0نادرست است ⟵ ولی ما آن را رد نمی‌کنیم.

مرحله نهم: محاسبه p-value

حالا به مهم‌ترین قسمت آزمون می‌رسیم.

از اطلاعات نمونه و روش آماری مناسب، یک p-value به دست می‌آوریم.

فرض کنید در این مثال:

p-value = 0.03

به دست آمده است.

p-valueرا نباید به‌عنوان احتمال درست بودن یا غلط بودن H_0 تفسیر کنیم.

p-valueدر واقع می‌گوید:

اگر فرض صفر درست باشد، مشاهده نتیجه‌ای به این اندازه یا شدیدتر چقدر محتمل است؟

در مثال ما، p-valueبرابر با ۰/۰۳ است.

حالا آن را با سطح معناداری مقایسه می‌کنیم:

p-value = 0.03

α = 0.05

چون:

۰.۰۳ < 0.05

است، نتیجه آزمون از نظر آماری معنادار است و ما فرض صفر را رد می‌کنیم.

مرحله دهم: معناداری آماری یعنی چه؟

اینجا باید یک نکته مهم را یاد بگیریم.

وقتی می‌گوییم نتیجه از نظر آماری معنادار است، منظورمان این نیست که حتماً این تفاوت از نظر عملی یا اقتصادی مهم است.

معناداری آماری یعنی داده‌های ما شواهد کافی برای رد فرض صفر فراهم کرده‌اند، با توجه به سطح معناداری انتخاب‌شده.

مثلا ممکن است میانگین عمر لامپ‌ها فقط ۱۰ ساعت با مقدار ادعایی شرکت تفاوت داشته باشد، اما به دلیل حجم نمونه بسیار بزرگ، این تفاوت از نظر آماری معنادار شود.

بنابراین در یک تحلیل مناسب، فقط به p-value نگاه نمی‌کنیم و بهتر است اندازه اثر، فاصله اطمینان و اهمیت عملی تفاوت را هم بررسی کنیم.

مرحله یازدهم: آزمونt

حالا سؤال این است که برای این مثال از چه آزمونی استفاده کنیم؟

فرض کنید انحراف معیار جامعه را نمی‌دانیم.

ما فقط اطلاعات ۳۰ لامپ نمونه را در اختیار داریم و انحراف معیار را از خود نمونه محاسبه می‌کنیم.

در چنین شرایطی برای آزمون میانگین، می‌توانیم از آزمون t یک‌نمونه‌ای یا One-Sample t-test استفاده کنیم.

در این آزمون، میانگین نمونه را با مقدار فرض‌شده، یعنی ۱۰۰۰ ساعت، مقایسه می‌کنیم.

اطلاعات اصلی ما عبارت‌اند از:

میانگین نمونه = ۹۷۰ ساعت

مقدار فرض‌شده = ۱۰۰۰ ساعت

حجم نمونه = ۳۰

و انحراف معیار نمونه که از داده‌های ۳۰ لامپ محاسبه شده است.

آزمون t با استفاده از این اطلاعات، میزان فاصله میانگین نمونه از مقدار فرض‌شده را نسبت به تغییرپذیری داده‌ها بررسی می‌کند و در نهایت یک آماره t و یک p-value در اختیار ما قرار می‌دهد.

مرحله دوازدهم: نتیجه‌گیری

حالا فرض کنید خروجی آزمون t به ما می‌گوید:

p-value = 0.03

و ما از ابتدا:

α = 0.05

را تعیین کرده‌ایم.

چون p-value کمتر از سطح معناداری است، فرض صفر را رد می‌کنیم.

بنابراین در سطح معناداری ۵ درصد، داده‌های نمونه شواهد آماری کافی برای حمایت از این نتیجه فراهم می‌کنند که میانگین عمر لامپ‌ها با ۱۰۰۰ ساعت تفاوت دارد.

دقت کنید که ما نمی‌گوییم:

»با احتمال ۹۷ درصد میانگین متفاوت است«.

چنین تفسیری از p-value درست نیست.

همچنین نمی‌گوییم:

H_0»با احتمال ۳ درصد درست است«.

این هم تفسیر نادرستی از p-value است.

ما فقط می‌گوییم:

اگر H_0 درست باشد، مشاهده چنین داده‌هایی یا داده‌هایی شدیدتر از آن، به اندازه‌ای کم‌احتمال است که در سطح معناداری ۵ درصد، H_0را رد می‌کنیم.

ارتباط فاصله اطمینان و آزمون فرض

حالا یک نکته بسیار جالب وجود دارد.

در مثال ما، فاصله اطمینان ۹۵ درصد را به صورت زیر در نظر گرفتیم:

۹۴۰ تا ۱۰۰۰ ساعت

و فرض صفر ما این بود:

H_0: μ = ۱۰۰۰

در یک آزمون دوطرفه با سطح معناداری ۵ درصد، ارتباط نزدیکی بین فاصله اطمینان ۹۵ درصد و آزمون فرض با α = 0.05 وجود دارد.

به‌طور کلی، اگر مقدار فرض‌شده خارج از فاصله اطمینان ۹۵ درصد باشد، با آزمون دوطرفه متناظر، H_0در سطح ۵ درصد رد می‌شود.

این ارتباط به دانشجو کمک می‌کند بفهمد که فاصله اطمینان و آزمون فرض دو مفهوم کاملاً جدا از هم نیستند؛ بلکه هر دو برای بررسی عدم‌قطعیت و شواهد آماری استفاده می‌شوند.

جمع‌بندی مثال

پس اجازه بدهید کل مثال را یک بار از ابتدا مرور کنیم.

شرکت ادعا می‌کند:

میانگین عمر لامپ‌ها = ۱۰۰۰ ساعت

ما نمی‌توانیم تمام لامپ‌ها را بررسی کنیم، بنابراین یک نمونه ۳۰تایی انتخاب می‌کنیم.

از نمونه، میانگین را محاسبه می‌کنیم:

$\bar{x} = ۹۷۰$ ساعت

این مقدار یک آماره است.

میانگین واقعی جامعه، یعنیμ ، یک پارامتر است که آن را نمی‌دانیم.

به دلیل نمونه‌گیری، احتمال وجود خطای نمونه‌گیری وجود دارد.

برای تخمین μ می‌توانیم یک فاصله اطمینان محاسبه کنیم.

سپس برای بررسی ادعای شرکت، فرض‌ها را تعریف می‌کنیم:

$H_0: \mu = ۱۰۰۰$

$H_1: \mu \neq ۱۰۰۰$

سطح معناداری را برابر ۰/۰۵ در نظر می‌گیریم.

با استفاده از آزمونt ، p-valueرا به دست می‌آوریم.

اگر:

$p\text{-value} < 0.05$

باشد، H_0را رد می‌کنیم.

در این مسیر باید توجه داشته باشیم که امکان خطای نوع اول و نوع دوم وجود دارد.

و در نهایت، نتیجه آماری را در کنار فاصله اطمینان و اهمیت عملی تفسیر می‌کنیم.

این دقیقاً همان کاری است که آمار استنباطی انجام می‌دهد:

از یک نمونه نسبتاً کوچک، اطلاعاتی درباره یک جامعه بزرگ‌تر به دست می‌آوریم و در عین حال، عدم‌قطعیت و احتمال خطا را نیز در تصمیم‌گیری آماری در نظر می‌گیریم.