

نسبت درستی

Likelihood Ratio

تهیه و تدوین توسط سایت: www.GeneProtocols.ir



مرجع دانلود کتابها و منابع علوم زیستی | قدرت گرفته از بیان

فهرست مطالب این مقاله

۲	تعریف مفهوم.....
۲	نحوه محاسبه نسبت درستی.....
۳	اهمیت محاسبه نسبت درستی.....
۴	مزیت استفاده از نسبت درستی نسبت به بررسی ساده فرض آماری.....
۵	یک مثال ساده.....
۶	یک مثال عملی.....
۸	مثالی در زمینه مدیریت.....
۹	مثالی در زمینه پزشکی.....
۱۰	استفاده از روش نسبت درستی برای داده های دارای توزیع غیر نرمال.....
۱۰	استفاده از روش نسبت درستی برای داده های پویا.....
۱۱	مثالی از استفاده از روش نسبت درستی در تحلیل داده های پویا.....
۱۲	منابع.....

تعریف مفهوم

نسبت درست‌نمایی یا "ضریب درست‌نمایی" (likelihood ratio) "در فرضیات آماری و تحلیل تصمیمی به کار می‌رود. این نسبت، نسبت احتمال اینکه داده‌های مشاهده شده مطابق با فرضیه A باشند به احتمال اینکه داده‌ها مطابق با فرضیه B باشند را محاسبه می‌کند.

به طور ریاضی، نسبت درست‌نمایی برابر است با:

$$LR = P(\text{Data} \mid \text{Hypothesis A}) / P(\text{Data} \mid \text{Hypothesis B})$$

در این فرمول، $P(\text{Data} \mid \text{Hypothesis A})$ نشان‌دهنده احتمال دیده شدن داده‌های مشاهده شده در صورتی است که فرضیه A درست باشد و $P(\text{Data} \mid \text{Hypothesis B})$ نشان‌دهنده احتمال دیده شدن داده‌های مشاهده شده در صورتی است که فرضیه B درست باشد.

نسبت درست‌نمایی می‌تواند به عنوان یک معیار برای تصمیم‌گیری در مورد درستی یا نادرستی فرضیات استفاده شود. اگر نسبت درست‌نمایی بزرگتر از ۱ باشد، احتمال بیشتری برای درستی فرضیه A نسبت به فرضیه B وجود دارد. اگر نسبت درست‌نمایی کوچکتر از ۱ باشد، احتمال بیشتری برای درستی فرضیه B نسبت به فرضیه A وجود دارد. اگر نسبت درست‌نمایی برابر با ۱ باشد، هیچ اطلاعاتی در مورد درستی یا نادرستی فرضیات به دست نمی‌آید.

نحوه محاسبه نسبت درست‌نمایی

برای محاسبه نسبت درست‌نمایی، باید ابتدا احتمال دیده شدن داده‌های مشاهده شده در صورتی که فرضیه A درست باشد ($P(\text{Data} \mid \text{Hypothesis A})$) و احتمال دیده شدن داده‌های مشاهده شده در صورتی که فرضیه B درست باشد ($P(\text{Data} \mid \text{Hypothesis B})$) را محاسبه کنید. سپس، با تقسیم احتمال دیده شدن داده‌های مشاهده شده در صورتی که فرضیه A درست باشد بر احتمال دیده شدن داده‌های مشاهده شده در صورتی که فرضیه B درست باشد، نسبت درست‌نمایی به دست می‌آید.

برای محاسبه احتمال دیده شدن داده‌های مشاهده شده در صورتی که فرضیه A یا B درست باشد، می‌توانید از روش‌های مختلفی مانند تحلیل آماری، تحلیل بیزی، مدل سازی آماری و ... استفاده کنید. برای مثال، در مدل ساده‌ای که از آن استفاده می‌شود، فرضیه A و B فقط دو فرضیه در نظر گرفته می‌شوند و برای محاسبه نسبت درست‌نمایی، احتمال دیده شدن داده‌های مشاهده شده در هر یک از این دو فرضیه، با استفاده از فرمول زیر محاسبه می‌شود:

$$LR = P(\text{Data} | \text{Hypothesis A}) / P(\text{Data} | \text{Hypothesis B})$$

در این فرمول، $P(\text{Data} | \text{Hypothesis A})$ و $P(\text{Data} | \text{Hypothesis B})$ به ترتیب نشان‌دهنده احتمال دیده شدن داده‌های مشاهده شده در صورتی که فرضیه A و B درست باشند هستند. سپس با تقسیم $P(\text{Data} | \text{Hypothesis A})$ بر $P(\text{Data} | \text{Hypothesis B})$ ، نسبت درست‌نمایی به دست می‌آید.

اهمیت محاسبه نسبت درست‌نمایی

اندازه‌گیری نسبت درست‌نمایی بسیار مهم است زیرا این مقدار می‌تواند به عنوان یک شاخص قطعی در تصمیم‌گیری‌های آماری و تحلیل تصمیمی مورد استفاده قرار گیرد. به طور کلی، نسبت درست‌نمایی بزرگتر از ۱ نشان می‌دهد که فرضیه A احتمالاً درست تر از فرضیه B است، در حالی که نسبت درست‌نمایی کوچکتر از ۱ نشان می‌دهد که فرضیه B احتمالاً درست تر از فرضیه A است. اگر نسبت درست‌نمایی برابر با ۱ باشد، هیچ اطلاعاتی به دست نمی‌آید و نمی‌توان تصمیمی قطعی در مورد درستی یا نادرستی فرضیات گرفت.

در صورتی که نسبت درست‌نمایی بسیار بزرگ باشد، به عنوان مثال بیشتر از ۱۰۰، این نتیجه را می‌دهد که فرضیه A بسیار احتمالی تر از فرضیه B است و این به عنوان یک شواهد قوی برای درستی فرضیه A در نظر گرفته می‌شود. به عنوان مثال، در مورد تشخیص بیماری‌های خطرناک، یک نسبت درست‌نمایی بسیار بزرگ بسیار قابل توجه است و می‌تواند به عنوان شواهدی قوی برای تشخیص بیماری استفاده شود.

به علاوه، با افزایش اندازه نمونه، نسبت درستنمایی به صورت معمول بهبود می‌یابد و به دلیل کاهش احتمال خطای نمونه‌گیری، این روش می‌تواند به دقت بیشتری در پیش‌بینی نتایج برسد. بنابراین، اندازه‌گیری نسبت درستنمایی می‌تواند به عنوان یک ابزار مفید در تحلیل داده‌های آماری و تصمیم‌گیری‌های مورد استفاده قرار گیرد.

مزیت استفاده از نسبت درستنمایی نسبت به بررسی ساده فرض آماری

استفاده از نسبت درستنمایی برای آزمون فرضیات آماری دارای مزیت‌هایی نسبت به روش‌های سنتی آماری مانند تست t و آزمون فرضیه‌های پارامتری است. مزیت‌های قابل اشاره را می‌توان در موارد ذیل خلاصه نمود:

- محاسبه نسبت درستنمایی برای تمام فرضیات ممکن، از جمله فرضیات آماری و غیرآماری، ممکن است به این دلیل که نسبت درستنمایی محاسبه شده بر اساس اطلاعات مشاهده شده است و نه بر اساس فرضیاتی که در آزمایش مورد بررسی قرار می‌گیرند. این مزیت می‌تواند در مواقعی که فرض آماری دقیق نیست و یا اطلاعات بیشتری در دسترس است، مفید باشد.
- نسبت درستنمایی برای فرضیات پویاست^۱ و با تغییر در داده‌های ورودی، مقدار آن تغییر می‌کند. بنابراین، این روش می‌تواند به عنوان یک ابزار بسیار قدرتمند در تحلیل داده‌های پویا و آزمایش فرضیات متغیر استفاده شود.
- در مقایسه با آزمایش فرضیه‌های پارامتری، نسبت درستنمایی به دلیل عدم نیاز به فرضیه‌های خاص برای پارامترها، از نظر محاسباتی ساده‌تر است.
- نسبت درستنمایی به عنوان یک شاخص عملکرد، برای مقایسه مدل‌های مختلف استفاده می‌شود و دارای تعبیر آسانی است.

^۱ فرضیات پویا در حوزه‌ی ریاضیات و علوم کامپیوتر به فرضیاتی گفته می‌شود که در آنها شرایط و قوانینی تعریف می‌شود که با گذشت زمان تغییر می‌کنند. به عبارتی، در فرضیات پویا، فرضیات و شرایطی که مورد بررسی قرار می‌گیرند، به صورت پویا و زمان‌بندی شده هستند و نمی‌توان آنها را به صورت ثابت در نظر گرفت. این فرضیات معمولاً در مدل‌های ریاضیاتی برای بررسی سیستم‌های پویا، مانند سیستم‌های خودرو، شبکه‌های ارتباطی، سیستم‌های بانکی و غیره استفاده می‌شوند. در این مدل‌ها، شرایط و فرضیاتی که با گذشت زمان تغییر می‌کنند، مانند وضعیت خودروها، تعداد کاربران در شبکه‌های ارتباطی و تغییرات در موجودی بانک، به صورت پویا مدل شده و برای پیش‌بینی رفتار سیستم در آینده استفاده می‌شوند.

• در مواردی که اندازه نمونه کوچک است و یا توزیع داده ها ناشناخته است، نسبت درستنمایی می‌تواند بهترین روش برای آزمون فرضیات باشد.

با توجه به مزایایی که برای نسبت درستنمایی بیان شد، می‌توان گفت که این روش به عنوان یک ابزار مفید در تحلیل داده‌های آماری و تصمیم‌گیری‌های مورد استفاده قرار می‌گیرد. با این حال، برای استفاده بهینه از این روش، نیاز به دانش کافی در زمینه آمار و تحلیل داده‌ها و تفسیر نتایج حاصل از آن دارید.

یک مثال ساده

در اینجا با ذکر یک مثال، محاسبه نسبت درستنمایی را توضیح می‌دهیم:

فرض کنید یک شرکت تولید کننده قرص ضد درد است و می‌خواهد بفهمد آیا قرص آنها موثرتر از قرص رقیبشان است یا خیر. برای این منظور، شرکت برای انجام آزمایش، ۲۰۰ نفر را به دو گروه تقسیم می‌کند. گروه اول ۱۰۰ نفری قرص شرکت تولید کننده و گروه دوم ۱۰۰ نفری قرص رقیب را دریافت می‌کنند. پس از یک ماه، تعدادی از اعضای هر گروه به شرکت بازخورد می‌دهند که آیا مصرف قرص برای آنها مفید بوده است یا خیر.

قرص شرکت تولید کننده موثرتر است: فرضیه A

قرص رقیب موثرتر است: فرضیه B

حال با توجه به داده های جمع آوری شده، احتمال دیده شدن این داده ها در صورتی که فرضیه A یا B درست باشد را محاسبه می‌کنیم. فرض کنید ۷۰ نفر از گروه شرکت تولید کننده و ۵۵ نفر از گروه رقیب بهبود یافته‌اند. در این صورت:

$$P(\text{Data} \mid \text{Hypothesis A}) =$$

احتمال اینکه ۷۰ نفر از گروه شرکت تولید کننده و ۵۵ نفر از گروه رقیب بهبود یافته باشند، در صورتی که قرص شرکت تولید کننده موثرتر باشد.

$$P(\text{Data} \mid \text{Hypothesis B}) =$$

احتمال اینکه ۷۰ نفر از گروه شرکت تولید کننده و ۵۵ نفر از گروه رقیب بهبود یافته باشند، در صورتی که قرص رقیب موثرتر باشد.

سپس با تقسیم احتمال دیده شدن داده های مشاهده شده در صورتی که فرضیه A درست باشد بر احتمال دیده شدن داده های مشاهده شده در صورتی که فرضیه B درست باشد، نسبت درست‌نمایی به دست می‌آید:

$$LR = P(\text{Data} \mid \text{Hypothesis A}) / P(\text{Data} \mid \text{Hypothesis B})$$

اگر احتمال دیده شدن داده های مشاهده شده در صورتی که فرضیه A درست باشد، مثلاً ۰/۸ و احتمال دیده شدن داده های مشاهده شده در صورتی که فرضیه B درست باشد، مثلاً ۰/۲ بود، نسبت درست‌نمایی برابر می‌شود با:

$$LR = 0.8 / 0.2 = 4$$

در این صورت، نسبت درست‌نمایی برابر با ۴ است. این بدان معناست که احتمال بیشتری برای درستی فرضیه A نسبت به فرضیه B وجود دارد و قابل قبول است که شرکت تولید کننده قرص موثرتری را تولید کرده باشد.

یک مثال عملی

فرض کنید یک شرکت موبایل، یک تبلت جدید را با نام "تبلت آ" معرفی کرده است و مدعی است که این تبلت بهترین تبلت موجود در بازار است. برای بررسی این مدعا، تعدادی از مصرف کنندگان آزمایشی را به صورت تصادفی انتخاب کرده و آنها را به دو گروه تقسیم می‌کند. گروه اول به عنوان گروه آزمایشی از تبلت "تبلت آ" استفاده می‌کنند و گروه دوم به عنوان گروه کنترل از یک تبلت دیگر استفاده می‌کنند. پس از یک دوره زمانی، نتایج پرسشنامه رضایتمندی مصرف کنندگان جمع‌آوری شده و نسبت درست‌نمایی برای مقایسه رضایتمندی مصرف کنندگان از دو تبلت محاسبه می‌شود.

تبلت آ" بهترین تبلت موجود در بازار است": فرضیه A
 تبلت آ" بهترین تبلت موجود در بازار نیست": فرضیه B

حال با توجه به داده‌های جمع‌آوری شده، نسبت درست‌نمایی محاسبه می‌شود. فرض کنید که نتایج آزمایش به شرح زیر باشد:

- گروه آزمایشی (تبلت آ): ۸۰٪ رضایتمندی
- گروه کنترل (تبلت ب): ۶۰٪ رضایتمندی

حال با محاسبه نسبت درست‌نمایی به دست می‌آوریم:

$$\text{نسبت درست‌نمایی} = \frac{P(\text{فرضیه A} \mid \text{داده های آزمایشی})}{P(\text{فرضیه B} \mid \text{داده های آزمایشی})}$$

با توجه به اطلاعاتی که در دست داریم، می‌توانیم مقدار

$$P(\text{فرضیه A} \mid \text{داده های آزمایشی})$$

9

$$P(\text{فرضیه B} \mid \text{داده های آزمایشی})$$

را محاسبه کنیم:

$$P(\text{فرضیه A} \mid \text{داده های آزمایشی}) = 0.8$$

$$P(\text{فرضیه B} \mid \text{داده های آزمایشی}) = 0.6$$

بنابراین، نسبت درست‌نمایی برابر است با:

$$\text{نسبت درست‌نمایی} = \frac{0.8}{0.6} = 1.33$$

با توجه به اینکه نسبت درستنمایی بزرگتر از ۱ است، نتیجه می‌گیریم که فرضیه A (یعنی تابلت آ بهترین تابلت موجود در بازار است) بهتر از فرضیه B (یعنی تابلت آ بهترین تابلت موجود در بازار نیست) در نظر گرفته می‌شود.

بنابراین، با توجه به نتایج آزمایش، این شرکت می‌تواند ادعای خود را مبنی بر معرفی تابلت آ به عنوان بهترین تابلت موجود در بازار، تایید کند.

مثالی در زمینه مدیریت

یکی از کاربردهای نسبت درستنمایی در مدیریت صنعتی، در بررسی کیفیت تولید محصولات است. فرض کنید یک شرکت تولیدی محصولات الکترونیکی یک سری تست کنترل کیفیت بر روی دستگاه‌های تولیدی خود انجام می‌دهد تا مطمئن شود که تولیدات مطابق با استانداردهای مورد نظر است.

در اینجا، فرضیه یک (H_1) این است که دستگاه‌های تولیدی به طور متوسط تنها ۵٪ محصولات با خطای نامطلوب تولید می‌کنند. فرضیه دو (H_2) این است که دستگاه‌های تولیدی به طور متوسط ۱۰٪ محصولات با خطای نامطلوب تولید می‌کنند.

حال، با انجام تست کنترل کیفیت بر روی ۱۰۰۰ دستگاه تولیدی، ۷۰ دستگاه با خطای نامطلوب تولید شده‌اند. برای ارزیابی این داده‌ها به کمک نسبت درستنمایی، می‌توان به شکل زیر عمل کرد:

$$L = P(\text{Data} | H_1) / P(\text{Data} | H_2)$$

در اینجا،

$$\text{Data} = 70$$

که تعداد دستگاه‌هایی است که با خطای نامطلوب تولید شده‌اند. برای محاسبه احتمال‌های مربوط به هر دو فرضیه، می‌توان از توزیع باینومیال استفاده کرد. برای فرضیه یک، می‌توان نوشت:

$$P(\text{Data} | H_1) = (1000 \text{ choose } 70) * (0.05)^{70} * (0.95)^{930}$$

و برای فرضیه دو، می‌توان نوشت:

$$P(\text{Data} | H_2) = (1000 \text{ choose } 70) * (0.1)^{70} * (0.9)^{930}$$

با جایگذاری این مقادیر در فرمول نسبت درست‌نمایی، می‌توان L را محاسبه کرد:

$$L = P(\text{Data} | H_1) / P(\text{Data} | H_2) = 0.0019$$

با توجه به اینکه L کمتر از ۱ است، می‌توان نتیجه گرفت که فرضیه دو (H2) بهترین فرضیه برای توصیف داده‌هاست. به عبارت دیگر، احتمال اینکه دستگاه‌های تولیدی به طور متوسط ۱۰٪ محصول با خطای نامطلوب تولید کنند، بیشتر از احتمال اینکه فرضیه یک درست باشد است.

مثالی در زمینه پزشکی

یکی از مثال‌های کاربردی نسبت درست‌نمایی در پزشکی، تعیین اثربخشی یک درمان برای کاهش خطر ابتلا به بیماری است. به عنوان مثال، یک گروه تحت پژوهش شامل ۱۰۰ نفر با خطر بالای ابتلا به بیماری قلبی عروقی را در نظر بگیرید. این گروه به دو زیرگروه تقسیم می‌شود، یک زیرگروه درمانی که یک درمان خاص دریافت می‌کنند و یک زیرگروه کنترل که هیچ درمانی دریافت نمی‌کنند. پس از یک دوره ۱۰ ساله، تعداد افرادی که در هر گروه به بیماری قلبی عروقی مبتلا شده‌اند، مشخص شده است. در گروه درمانی، ۱۰ نفر به بیماری قلبی عروقی مبتلا شده‌اند در حالی که در گروه کنترل، ۲۰ نفر به بیماری قلبی عروقی مبتلا شده‌اند.

برای بررسی تأثیر درمان بر کاهش خطر ابتلا به بیماری قلبی عروقی، از نسبت درست‌نمایی استفاده می‌کنیم. با استفاده از روابط زیر:

$$\text{نسبت درست‌نمایی} = (a/b) / (c/d)$$

در اینجا، a تعداد افراد مبتلا به بیماری قلبی عروقی در گروه درمانی، b تعداد افراد سالم در گروه درمانی، c تعداد افراد مبتلا به بیماری قلبی عروقی در گروه کنترل و d تعداد افراد سالم در گروه کنترل است.

با توجه به این مقادیر، می‌توان نسبت درستنمایی را به شکل زیر محاسبه کرد:

$$\text{نسبت درستنمایی} = \frac{۱۰ / ۹۰}{۲۰ / ۸۰} = ۰/۴۴$$

با توجه به اینکه مقدار نسبت درستنمایی کمتر از ۱ است، می‌توان نتیجه گرفت که درمان مورد بررسی تأثیر چشمگیری در کاهش خطر ابتلا به بیماری قلبی عروقی ندارد. به عبارت دیگر، احتمال ابتلا به بیماری قلبی عروقی در گروه درمانی نسبت به گروه کنترل کمتر است، اما این تفاوت به اندازه‌ای نیست که بتواند به عنوان یک تفاوت معنادار در نظر گرفته شود.

استفاده از روش نسبت درستنمایی برای داده‌های دارای توزیع غیر نرمال

روش نسبت درستنمایی به نوع توزیع داده‌ها وابسته نیست و در مواردی که داده‌ها توزیع نرمال ندارند نیز قابل استفاده است. به طور کلی، توزیع نرمال تنها یکی از توزیع‌های موجود است و در بسیاری از موارد، داده‌ها به توزیع نرمال نمی‌آیند. در چنین شرایطی، از روش‌های آماری غیرپارامتری مانند آزمون ویلکاکسون یا آزمون کروسکال-والیس برای بررسی تفاوت بین دو یا چند گروه داده‌ای استفاده می‌شود. با این حال، در صورتی که هدف ما بررسی تأثیر یک متغیر مستقل بر یک متغیر وابسته است، می‌توان از روش نسبت درستنمایی استفاده کرد و نتایج قابل قبولی را به دست آورد.

استفاده از روش نسبت درستنمایی برای داده‌های پویا

روش نسبت درستنمایی در حقیقت یک روش برای بررسی تأثیر یک متغیر مستقل بر متغیر وابسته است و تنها به داده‌های موجود در دو گروه (مثلاً گروه کنترل و گروه آزمایشی) نیاز دارد. بنابراین، در مواردی که داده‌ها پویا باشند و به صورت زمانی جمع‌آوری شوند نیز می‌توان از روش نسبت درستنمایی استفاده کرد.

در واقع، در این حالت می‌توان برای هر زمان مشاهده شده، داده‌های مربوط به هر گروه را جداگانه بررسی کرد و نسبت درستنمایی را برای هر زمان محاسبه کرد. با این روش، می‌توان تغییرات زمانی در تأثیر متغیر مستقل بر متغیر وابسته را بررسی کرد و به نتایج دقیق‌تری رسید.

بنابراین، روش نسبت درست‌نمایی در داده‌های پویا نیز به راحتی قابل استفاده است و می‌تواند نتایج دقیقی را برای بررسی تأثیر یک متغیر مستقل بر متغیر وابسته در طول زمان برای هر گروه داده‌ای به دست آورد. با این حال، همانند هر روش آماری دیگر، باید به دقت داده‌ها را بررسی کرد و از مدل مناسب برای هر نوع داده استفاده کرد.

مثالی از استفاده از روش نسبت درست‌نمایی در تحلیل داده‌های پویا

فرض کنید شما در حال بررسی اثر یک تبلیغات تلویزیونی بر فروش یک محصول هستید. برای این منظور، دو گروه کنترل و آزمایش را تعریف می‌کنید. در گروه کنترل، تبلیغاتی نشان داده نمی‌شود، اما در گروه آزمایش، تبلیغات نمایش داده می‌شود.

برای بررسی تأثیر تبلیغات، شما تعداد فروش محصول را در هر دو گروه قبل و بعد از اجرای تبلیغات جمع‌آوری کرده‌اید. برای تحلیل این داده‌ها، می‌توانید از روش نسبت درست‌نمایی استفاده کنید.

فرض کنید که در گروه کنترل، ۱۰۰۰ محصول فروخته شده است و در گروه آزمایش، ۱۲۰۰ محصول فروخته شده است. با استفاده از فرمول زیر، می‌توانید نسبت درست‌نمایی بین دو گروه را محاسبه کنید:

$$\lambda = (x_2 / x_1)^{(1 / (t_2 - t_1))}$$

در اینجا، x_1 تعداد فروش در گروه کنترل و x_2 تعداد فروش در گروه آزمایش است. t_1 و t_2 نیز زمان‌های مشابه برای دو گروه هستند. با استفاده از فرمول، می‌توان نسبت درست‌نمایی بین دو گروه را محاسبه کرد. برای مثال، در اینجا:

$$\lambda = (1200 / 1000)^{(1 / (2 - 1))} = 1.2$$

با توجه به اینکه مقدار λ بیشتر از ۱ است، می‌توان نتیجه گرفت که تبلیغات تلویزیونی تأثیر مثبتی بر فروش محصول داشته است.

- Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences* (4th ed.). Prentice Hall.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Chapman & Hall/CRC.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773-795.
- Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14(5), 779-804.