

تفکر بیزی داشته باش

آمار بیزی ساده شد!

نویسنده: آلن. ب. داوونی

مترجم: امیرعباس اسدی

فهرست مطالب

۵	۱ قضیه‌ی بیز
۵	۱.۱ احتمال شرطی
۶	۲.۱ احتمال توأم
۶	۳.۱ مسئله‌ی کوکی
۶	۴.۱ قضیه‌ی بیز
۸	۵.۱ تفسیر در طول زمان
۸	۶.۱ مسئله‌ی M & M
۸	۷.۱ مسئله‌ی مونتی هال
۸	۸.۱ بحث

فصل ۱

قضیه‌ی بیز

۱.۱ احتمال شرطی

ایده‌ی بنیادی پشت تمام آمار بیزی قضیه‌ی بیز^۱ است، قضیه‌ی ای که به طرز شگفت‌انگیزی بدست آوردنش آسان بوده و همچنین امکان درک احتمال شرطی را فراهم می‌کند. بحث را از احتمال آغاز کرده، سپس احتمال شرطی، قضیه‌ی بیز و در آخر آمار بیزی را بررسی می‌کنیم.

احتمال عددی بین 0 و 1 (شامل این دو هم می‌شود) و نشانگر درجه‌ی باور ما به یک حقیقت یا پیشبینی است. مقدار 1 نشان دهنده اطمینان از درستی یک حقیقت و یا اطمینان از به وقوع پیوستن یک پیشبینی است. مقدار 0 هم نشانگر نادرستی یک حقیقت است.

مقادیر میانی نشان دهنده‌ی درجه‌ی اطمینان و قطعیت هستند. مقدار 0.5 که اغلب به صورت 50% نوشته می‌شود، به این معناست که وقوع یک پیشبینی به اندازه‌ی عدم وقوع آن محتمل است. مثلاً احتمال اینکه یک سکه‌ی پرتاب شده شیر باشد بسیار نزدیک 50% است.

یک احتمال شرطی^۲ احتمالی بر اساس برخی از اطلاعات قبلی است. برای مثال من می‌خواهم بدانم احتمال اینکه در سال آینده دچار یک حمله‌ی قلبی شوم چقدر است. بر طبق CDC، سالانه حدود 785,000 آمریکایی اولین حمله‌ی قلبی خود را تجربه می‌کنند.^۳ جمعیت ایالات متحده حدوداً 311 میلیون نفر است، بنابراین احتمال اینکه یک آمریکایی که به طور تصادفی انتخاب شده در سال آینده دچار یک حمله‌ی قلبی شود، حدوداً 0.3% است. اما من که یک آمریکایی که به صورت تصادفی انتخاب شده باشد نیستم، همه‌گیرشناسان پی برده‌اند که عوامل زیادی بر میزان خطر وقوع حمله‌ی قلبی تأثیر می‌گذارند؛ بر اساس آن عوامل، احتمال اتفاق افتادن یک حمله‌ی قلبی برای من می‌تواند بالاتر یا پایین‌تر از میانگین باشد. من یک مرد هستم، 45 سال سن و کلسترول تقریباً بالایی دارم. این عوامل شانس ابتلاء را افزایش می‌دهند. در حالیکه، فشار خون من پایین است، سیگار نمی‌کشم و این عوامل شانس را کاهش می‌دهند. من با وارد کردن اطلاعاتم در یک محاسبه‌کننده‌ی آنلاین به آدرس: <http://cvdrisk.nhlbi.nih.gov/calculator.asp> متوجه شدم احتمال اینکه سال آینده یک حمله‌ی قلبی را تجربه کنم چیزی حدود 0.2% و کمتر از میانگین است. این مقدار یک احتمال شرطی است چون بر اساس عواملی است که شرایط من محسوب می‌شوند.

معمولاً احتمال شرطی را به صورت $p(A|B)$ نمایش می‌دهند که معنای آن احتمال درستی A با فرض درستی B است. در این مثال، A نشان

¹Bayes's Theorem

²Conditional Probability

³<http://www.cdc.gov/heartdisease/facts.htm>

دهنده‌ی ابتلای من به حمله‌ی قلبی در سال آینده و B نشان دهنده‌ی شرایطی که نام بردم است.

۲.۱ احتمال توأم

احتمال توأم^۴ روشی برای بیان احتمال درستی دو چیز است. منظورمان از $p(A \text{ and } B)$ ، احتمال درستی هر دو رخداد A و B است. اگر شما در مورد احتمال پرتاب سکه و تاس بدانید احتمالا فرمول زیر را آموخته اید.

$$p(A \text{ and } B) = p(A)p(B) \quad \text{هشدار: این فرمول همیشه درست نیست}$$

برای مثال اگر دو سکه را پرتاب کنیم و معنی A شیر بودن سکه اول و معنی B شیر بودن سکه دوم باشد، خواهیم داشت $p(A) = p(B) = 0.5$ و بنابراین $p(A \text{ and } B) = p(A)p(B) = 0.25$. اما این فرمول صرفا به این دلیل درست است که A و B مستقل هستند. به عبارت دیگر دانستن نتیجه‌ی رخداد اول، احتمال رخداد دوم را تغییر نمی دهد، مطابق تعریف $p(B|A) = p(B)$. اکنون مثالی متفاوت که در آن A و B مستقل نیستند را بررسی می کنیم. فرض کنید A یعنی امروز باران می آید و B یعنی اینکه فردا باران می آید. اگر بدانیم که امروز باران آمده، امکان اینکه فردا هم باران ببارد بیشتر است بنابراین $p(B|A) > p(B)$ به صورت کلی احتمال توأم برابر است با

$$p(A \text{ and } B) = p(A)p(B|A)$$

برای هر دو رخداد A و B . با این اوصاف اگر احتمال باران آمدن در هر روز 0.5 باشد، احتمال باران آمدن در دو روز پشت سر هم 0.25 نبوده و احتمالا کمی بیشتر است.

۳.۱ مسئله‌ی کوکی

به زودی به سراغ قضیه‌ی بیز خواهیم رفت اما قبل از آن می خواهم اهمیت این قضیه را با یک مثال به نام مسئله‌ی کوکی روشن کنم^۵. فرض کنید دو کاسه پر از کوکی داریم. کاسه‌ی اول حاوی 30 کوکی وانیلی و 10 کوکی شکلاتی و کاسه‌ی دوم حاوی 20 عدد از هر کدام است. اکنون فرض کنید که یکی از کاسه ها را به صورت تصادفی انتخاب و بدون نگاه کردن در آن، یک از کوکی هایش را برمی داریم. اگر کوکی وانیلی باشد، چقدر احتمال دارد که از کاسه‌ی اول انتخاب شده باشد؟ این یک احتمال شرطی است؛ میخواهیم (وانیلی|کاسه‌ی اول) p را محاسبه کنیم که روش محاسبه اش چندان واضح به نظر نمی رسد. اما اگر به جای آن قرار بود احتمال وانیلی بودن کوکی را با فرض اینکه از کاسه‌ی اول انتخاب شده باشد محاسبه کنیم، سوال راحتی بود:

$$p(\text{کاسه‌ی اول}|\text{وانیلی}) = \frac{3}{4}$$

متأسفانه، $p(A|B)$ با $p(B|A)$ برابر نیست. اما راهی وجود دارد که با استفاده از یکی، دیگری را بدست آوریم: قضیه‌ی بیز.

۴.۱ قضیه‌ی بیز

اکنون هرآنچه که برای بدست آوردن قضیه‌ی بیز نیاز داریم را در اختیار داریم. بحث را از این حقیقت که ترکیب عطفی جابجایی پذیر است آغاز می کنیم؛ یعنی:

$$p(A \text{ and } B) = p(B \text{ and } A)$$

⁴Conjoint Probability

^۵بر اساس مثالی در http://en.wikipedia.org/wiki/Bayes'_theorem، که دیگر در این صفحه وجود ندارد

برای هر دو رخداد A و B .

سپس مقدار احتمال توأم را جایگذاری می‌کنیم:

$$p(A \text{ and } B) = p(A)p(B|A)$$

از آنجایی که معنای خاصی برای A و B در نظر نگرفته ایم، می‌توانیم آن‌ها را جابجا کنیم. با جابجایی آن‌ها خواهیم داشت:

$$p(B \text{ and } A) = p(B)p(A|B)$$

این تمام چیزی بود که نیاز داریم. اکنون اگر این‌ها را کنار هم بگذاریم، بدست می‌آید:

$$p(B)p(A|B) = p(A)p(B|A)$$

این یعنی دو راه برای محاسبه‌ی احتمال توأم داریم. اگر مقدار $p(A)$ را داشته باشید، می‌توانید آن را در احتمال شرطی $p(B|A)$ ضرب کنید. و یا اینکه می‌توانید از ضرب $p(B)$ در $p(A|B)$ برای محاسبه استفاده کنید. هر کدام را که انتخاب کنید، نتیجه یکسان است. در آخر طرفین را بر $p(B)$ تقسیم می‌کنیم:

$$p(A|B) = \frac{p(A)p(B|A)}{p(B)}$$

و این قضیه‌ی بیز است! شاید در نگاه اول به نظر نرسد، اما به طرز شگفت‌انگیزی قدرتمند است.

برای مثال می‌توانیم با استفاده از آن مسئله‌ی کوکی را حل کنیم. از B_1 برای نشان دادن رخداد بیرون آمدن کوکی از کاسه‌ی اول و از V برای نشان دادن وانیلی بودن کوکی استفاده می‌کنیم. با استفاده از قضیه‌ی بیز خواهیم داشت:

$$p(B_1|V) = \frac{p(B_1)p(V|B_1)}{p(V)}$$

عبارت سمت چپ همان چیزی است که به دنبالش بودیم: احتمال اینکه کوکی از کاسه‌ی اول انتخاب شده باشد، اگر بدانیم طعم آن وانیلی است. سمت راست عبارت است از:

- $p(B_1)$: احتمال اینکه بدون در نظر گرفتن طعم کوکی، کاسه‌ی اول را انتخاب کنیم. از آنجایی که طبق مسئله کاسه به صورت تصادفی انتخاب می‌شود، می‌توانیم فرض کنیم $p(B_1) = \frac{1}{2}$.
- $p(V|B_1)$: یعنی احتمال انتخاب یک کوکی وانیلی از کاسه‌ی اول که برابر است با $\frac{3}{4}$.

- $p(V)$: یعنی احتمال انتخاب کوکی وانیلی از هر دو کاسه. از آن جایی که شانس انتخاب کاسه‌ها یکسان است و کاسه‌ها حاوی تعداد یکسانی کوکی هستند، شانس انتخاب هر کوکی یکسان است. در دو کاسه مجموعاً 50 کوکی وانیلی و 30 کوکی شکلاتی است، پس $p(V) = \frac{5}{8}$.

با جایگذاری در قضیه‌ی بیز بدست می‌آید:

$$p(B_1|V) = \frac{\frac{1}{2} \times \frac{3}{4}}{\frac{5}{8}}$$

که با $\frac{3}{5}$ برابر است. پس انتخاب کوکی وانیلی برهانی^۶ برای کاسه‌ی اول است، چون بیرون آمدن کوکی‌های وانیلی از کاسه‌ی اول محتمل‌تر است. این مثال یکی از استفاده‌های قضیه‌ی بیز را نشان می‌دهد: فراهم کردن روشی برای بدست آوردن $p(A|B)$ از $p(B|A)$. این استراتژی در مواردی مانند مسئله‌ی کوکی، که محاسبه‌ی سمت راست قضیه‌ی بیز آسان‌تر از سمت چپ آن است، به کار می‌آید.

⁶evidence

۵.۱ تفسیر در طول زمان

می‌توان به قضیه‌ی بیز از زاویه‌ی دیگری هم نگاه کرد: این قضیه راهی برای تصحیح و بروزرسانی احتمال یک فرض مثل H را در صورت مشاهده‌ی داده‌هایی جدید مانند D فراهم می‌کند. این شیوه‌ی فکر کردن به قضیه‌ی بیز تفسیر در طول زمان^۷ نامیده می‌شود. احتمال فرضیات در طول زمان مشاهده‌ی داده‌های جدید تغییر می‌کند. اگر قضیه‌ی بیز را با H و D بازنویسی کنیم، خواهیم داشت:

$$p(H|D) = \frac{p(D|H)p(H)}{p(D)}$$

در این تفسیر هر کدام از عبارات بالا نامی دارند:

- $p(H)$ احتمال فرض پیش از مشاهده‌ی داده‌ها است که احتمال پیشین با به صورت مختصر پیشین^۸ نامیده می‌شود.
- $p(H|D)$ چیزی است که می‌خواهیم محاسبه کنیم، احتمال فرضیات پس از مشاهده‌ی داده‌ها است که پسین^۹ نامیده می‌شود.
- $p(D|H)$ احتمال تولید داده‌ها تحت فرض است که درست نمایی^{۱۰} نامیده می‌شود.
- $p(D)$ احتمال تولید داده‌ها تحت هر فرض ممکن است که ثابت نرمال سازی^{۱۱} نامیده می‌شود.

گاهی اوقات می‌توانیم احتمال پیشین را بر اساس اطلاعات پیشفرض خود محاسبه کنیم. مثلاً در مسئله‌ی کوکی مشخص شده است که احتمال انتخاب کاسه‌ها یکسان است. در دیگر موارد ممکن است پیشین به صورت ذهنی انتخاب شود و افراد به دلیل اطلاعات پیشفرض متفاوت و یا تفسیر متفاوت از اطلاعات یکسان، بر سر شیوه‌ی محاسبه‌ی آن دچار اختلاف شوند. محاسبه‌ی درست نمایی معمولاً آسان‌ترین قسمت کار است. در مسئله‌ی کوکی، اگر بدانیم کوکی از کدام کاسه انتخاب شده، می‌توانیم احتمال وانیلی بودن کوکی را محاسبه کنیم. محاسبه‌ی ثابت نرمال سازی گاهی نیاز به حقه دارد. این مقدار به عنوان احتمال مشاهده‌ی داده تحت هر فرض ممکن تعریف می‌شود اما در موارد کلی تعیین اینکه دقیقاً چه معنایی دارد، دشوار است. در اغلب موارد با متمایز کردن مجموعه‌هایی از فرضیات که ویژگی‌های خاصی دارند، ساده سازی انجام می‌دهیم. این دو ویژگی عبارتند از:

دو به دو ناسازگار^{۱۲}: حداقل یکی از فرض‌های مجموعه می‌تواند درست باشد و

کاملاً در برگیرنده^{۱۳}: حداقل یکی از فرضیات باید درست باشد.

در ادامه برای مجموعه‌هایی از فرضیات که دارای این ویژگی‌ها هستند کلمه‌ی مناسب را استفاده می‌کنیم. در مسئله‌ی کوکی تنها دو فرض داشتیم - کوکی از کاسه اول یا دوم بیرون آمده است - که این دو فرض دو به دو ناسازگار و کاملاً در برگیرنده هستند.

۶.۱ مسئله‌ی M & M

۷.۱ مسئله‌ی مونتی هال

۸.۱ بحث

⁷The Diachronic Interpretation

⁸prior

⁹posterior

¹⁰likelihood

¹¹normalizing constant

¹²Mutually exclusive

¹³Collectively exhaustive