

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
۷	۴.۱ قضیه‌ی بیز
۷	۵.۱ تفسیر در طول زمان
۷	۶.۱ مسئله‌ی M & M
۷	۷.۱ مسئله‌ی مونتی هال
۷	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 ۰ و ۱ (شامل این دو هم می‌شود) و نشانگر درجه‌ی باور ما به یک حقیقت یا پیشبینی است. مقدار ۱ نشان دهنده اطمینان از درستی یک حقیقت و یا اطمینان از به وقوع پیوستن یک پیشبینی است. مقدار ۰ هم نشانگر نادرستی یک حقیقت است.

مقادیر میانی نشان دهنده‌ی درجه‌ی اطمینان و قطعیت هستند. مقدار ۰.۵ که اغلب به صورت ۵۰٪ 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 ۵۰٪ 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 ۷۸۵,۰۰۰ آمریکایی اولین حمله‌ی قلبی خود را تجربه می‌کنند.^۳ جمعیت ایالات متحده حدوداً ۳۱۱ میلیون نفر است، بنابراین احتمال اینکه یک آمریکایی که به طور تصادفی انتخاب شده در سال آینده دچار یک حمله‌ی قلبی شود، حدوداً ۰.۳٪ 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، ۴۵ 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 ۰.۲٪ 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 احتمالاً 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 صرفاً به این دلیل درست است که A و B مستقل هستند. به عبارت دیگر دانستن نتیجه‌ی رخداد اول، احتمال رخداد دوم را تغییر نمی‌دهد، مطابق تعریف $p(B|A) = p(B)$. اکنون مثالی متفاوت که در آن A و B مستقل نیستند را بررسی می‌کنیم. فرض کنید A یعنی امروز باران می‌آید و B یعنی اینکه فردا باران می‌آید. 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 احتمالاً کمی بیشتر است.

۳.۱ مسئله‌ی کوکی

به زودی به سراغ قضیه‌ی بیز خواهیم رفت اما قبل از آن می‌خواهم اهمیت این قضیه را با یک مثال به نام مسئله‌ی کوکی روشن کنم. 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 کاسه‌ها را به صورت تصادفی انتخاب و بدون نگاه کردن در آن، یک از کوکی‌هایش را برمی‌داریم. 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 محاسبه‌اش چندان واضح به نظر نمی‌رسد. 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.

⁴Conjoint Probability

۴.۱ قضیه‌ی بیز

۵.۱ تفسیر در طول زمان

۶.۱ مسئله‌ی **M & M**

۷.۱ مسئله‌ی مونتی هال

۸.۱ بحث