

حذف ناممکن‌ها: از قصه تا واقعیت

کیوان آقابابائی سامانی *

دانشکده فیزیک، دانشگاه صنعتی اصفهان

۹ اردیبهشت ۱۴۰۴

شرلوك هولمز در كتاب *نشانه چهار روشی* را برای كشف حقیقت به كار می‌برد كه می‌توان اسمش را گذاشت روش حذف ناممکن‌ها: وقتی همه حالت‌های ناممکن را كنار گذاشته باشی، آنچه باقی می‌ماند، هرچه قدر هم نامحتمل، باید حقیقت باشد^۱.

هولمز در حالی كه سرش را تكان می‌داد گفت:

– تو به توصیه من عمل نمی‌كنی. چند بار به تو گفته‌ام كه وقتی ناممکن را حذف کرده باشی، آنچه باقی می‌ماند، هر قدر هم بعید، باید حقیقت باشد؟ می‌دانیم كه او از دریا از پنجره یا دودكش وارد نشده. این را هم می‌دانیم كه نمی‌توانسته در اتاق پنهان شده باشد، چون مخفیگاهی وجود ندارد. پس از كجا آمده؟

من فریاد زدم:

– از سوراخ سقف آمده؟

– معلوم است كه از آنجا آمده.^۲

این روش، مثل بسیاری از روش‌های دیگری كه هولمز به كار می‌گیرد، جذاب و هیجان‌انگیز است ولی آیا در عمل و در شرایط واقعی هم می‌توان چنین روش‌هایی را به همان سادگی

k1samani@gmail.com*

https://t.me/k1samani_channel

^۱ When you have eliminated the impossible, whatever remains, however improbable, must be the truth.

^۲ آرتور کانن دوویل، *نشانه چهار*، ترجمه مزده دقیقی، (انتشارات هرمس ۱۳۷۸).

به کار برد؟ واقعیت این است که در عمل ممکن است موانع فراوانی کاربرپذیری این روش را به چالش بکشد. در این یادداشت به دو مورد از این موانع نگاه دقیق‌تری می‌اندازیم. نخست این که تشخیص ناممکن بودن بعضی حالت‌ها که در ابتدا محتمل بوده‌اند با چه دقتی انجام می‌شود؟ آیا ممکن نیست خطایی در این تشخیص وجود داشته باشد؟ مثلاً در همین مکالمه که از کتاب نشانه چهار نقل شد هولمز به واتسن می‌گوید: «می‌دانیم که او از در یا از پنجره یا دودکش وارد نشده.» این «می‌دانیم» چه قدر دقیق است؟ آیا ممکن است خطایی در مشاهده یا جمع‌آوری شواهد وجود داشته باشد که این نتیجه‌گیری را نادقیق کند؟ دوم این که آیا همه حالت‌های ممکن از ابتدا در نظر گرفته شده‌اند؟ مثلاً آیا ممکن است که به جز در، پنجره، دودکش و سوراخ سقف راه دیگری هم برای ورود به اتاق بوده باشد که از نظر کارآگاه دور مانده باشد؟ چنین اتفاقاتی تا چه اندازه می‌تواند اعتبار نتیجه‌گیری نهایی را به خطر بیندازد؟

خطا در تشخیص ناممکن بودن حالت‌ها

برای بررسی خطای آزمون‌هایی که حالت‌های ناممکن را مشخص می‌کنند مثال ساده‌ای در نظر می‌گیریم. فرض کنید $n+1$ جعبه داریم که در یکی از آن‌ها یک توپ وجود دارد. جعبه‌ها را با $1, 2, \dots, n, z$ برچسب می‌زنیم هیچ اطلاعات اضافه دیگری که بتواند راهنمایی برای تشخیص جعبه محتوی توپ باشد نداریم، بنابراین فرض می‌کنیم که توپ می‌تواند با احتمال یکسان در هر یک از این $n+1$ جعبه باشد.^۳

فرض کنید A_i پیشامد قرار داشتن توپ در جعبه شماره i باشد. همچنین فرض کنید برای تشخیص این که یک جعبه تویی در درونش ندارد از آزمایشی مانند تکان دادن جعبه یا اسکن کردن آن با پرتو ایکس استفاده کنیم. پیشامدی را که در آن نتیجه آزمایش روی جعبه شماره i خالی بودن آن را نشان می‌دهد E_i می‌نامیم. این آزمایش ممکن است خطا

^۳ می‌توان مسئله را به شکل عام‌تری هم طرح کرد، مثلاً می‌توان فرض کرد که احتمال این که توپ در جعبه شماره i باشد p_i است و $\sum_{i=1}^n p_i + p_z = 1$.

داشته باشد، به این معنی که ممکن است توپ در جعبه شماره i باشد ولی نتیجه آزمایش خلاف این را گزارش کند. احتمال چنین خطایی را با r نشان می‌دهیم. به زبان ریاضی $P(E_i|A_i) = r$.

جعبه‌های شماره 1 تا n را آزمایش می‌کنیم و نتیجه آزمایش این است که توپ در هیچ‌یک از این جعبه‌ها نیست. طبق روش هولمز می‌توانیم بگوییم که توپ قطعاً در جعبه z است. اما احتمال خطا در آزمایش‌های ما وجود دارد و بنابراین ممکن است توپ در یکی از جعبه‌هایی باشد که آزمایش خالی بودن آن را نشان داده. در چنین شرایطی نمی‌توان گفت که توپ قطعاً در جعبه z است. سؤال درست این است که احتمال قرار داشتن توپ در جعبه z چه قدر است. چنین احتمالی با عبارت ریاضی زیر بیان می‌شود

$$P(A_z|E_1 \cap E_2 \cap \dots \cap E_n).$$

برای سادگی این احتمال را P می‌نامیم و برای محاسبه آن از قاعدهٔ بیز^۴ استفاده می‌کنیم

$$\begin{aligned} P &:= P(A_z|E_1 \cap E_2 \cap \dots \cap E_n) = \frac{P(A_z \cap E_1 \cap E_2 \cap \dots \cap E_n)}{P(E_1 \cap E_2 \cap \dots \cap E_n)}, \\ &= \frac{P(E_1 \cap E_2 \cap \dots \cap E_n|A_z)P(A_z)}{P(E_1 \cap E_2 \cap \dots \cap E_n|A_z)P(A_z) + \sum_{i=1}^n P(E_1 \cap E_2 \cap \dots \cap E_n|A_i)P(A_i)}. \end{aligned}$$

فرض می‌کنیم آزمایش‌های جعبه‌های مختلف مستقل باشند یعنی آزمایش یک جعبه روی نتیجهٔ آزمایش یک جعبهٔ دیگر اثری نداشته باشد. در این صورت خواهیم داشت

$$P(E_1 \cap E_2 \cap \dots \cap E_n|A_z) = (1 - r)^n$$

زیرا اگر توپ در جعبهٔ شماره z باشد یعنی همهٔ آزمایش‌ها نتیجهٔ درست داده‌اند و احتمال

⁴Bayes' rule

درست بودن نتیجه هر آزمایش $1 - r$ است. به همین ترتیب به سادگی می توان دید که

$$P(E_1 \cap E_2 \cap \dots \cap E_n | A_i) = r(1 - r)^{n-1}, i = 1, 2, \dots, n$$

و بنابراین

$$\begin{aligned} P &= \frac{(1 - r)P(A_z)}{r(1 - P(A_z)) + (1 - r)P(A_z)} \\ &= \frac{1 - r}{nr + 1 - r} \end{aligned}$$

که در آن از $P(A_z) = \frac{1}{n+1}$ استفاده کرده ایم. بیایید نگاهی به نتیجه این رابطه برای یک حالت مشخص بیندازیم. فرض کنید 10 جعبه داریم (یعنی $n = 9$) و آزمایش ما برای تشخیص خالی بودن جعبه ها 10 درصد خطا دارد (به این معنی که به طور میانگین از هر ده آزمایش یکی نتیجه نادرست می دهد). در این صورت رابطه بالا می گوید که $P = 0.5$. یعنی احتمال این که توپ در جعبه آخر (جعبه ای که آزمایش نشده) باشد 50 درصد است. به طور کلی اگر احتمال خطا در آزمایش تشخیص حالت های ناممکن برابر با احتمال همان حالتی باشد که در آخر و پس از حذف ناممکن ها قرار است به عنوان «حقیقت» معرفی شود، خطای تشخیص حقیقت 50 درصد خواهد بود!

خطا در تعیین همه حالت های ممکن

ایراد دیگری که می تواند کارایی این روش را به چالش بکشد این است که از ابتدا همه حالت های ممکن را تشخیص نداده باشیم و بعضی از آن ها از چشم ما دور مانده باشند. اگر به مثال بخش قبل برگردیم می توانیم فرض کنیم که مثلاً یک جعبه دیگر (جعبه شماره m وجود دارد) که در میان $n + 1$ جعبه موجود نیست و مثلاً پشت یک پرده مخفی شده است ولی توپ می تواند درون آن جعبه هم باشد. احتمال وجود توپ در آن جعبه را با

$P(A_m)$ نشان می‌دهیم. به عبارت دیگر اگرچه ما تصور می‌کنیم که حاصل جمع احتمال‌های وجود توپ در $n + 1$ جعبه موجود برابر با یک است ولی درواقع این احتمال کوچک‌تر از یک است: $\sum_{i=1}^n P(A_i) + P(A_z) = 1 - P(A_m)$.

بنابراین حتی اگر آزمایش‌های ما بی‌خطا باشند و بگویند که توپ در جعبه‌های شماره 1 تا n نیست باز هم ممکن است جعبه شماره z را باز کنیم و ببینیم که خالی است. احتمال چنین نتیجه ناگواری $P(A_m)$ است. برای اجتناب از مواجه شدن با چنین وضعیتی باید حداکثر تلاش را برای تشخیص و به حساب آوردن همه حالت‌های ممکن به خرج داد.

به عنوان آخرین مثال حالتی را در نظر می‌گیریم که هم آزمایش‌ها احياناً خطا داشته باشند و هم از ابتدا همه حالت‌های ممکن تعیین نشده باشند و مثلاً جعبه شماره m از قلم افتاده باشد. در این صورت به سادگی می‌توان دید که وقتی آزمایش‌ها نشان می‌دهند که توپ در هیچ‌یک از جعبه‌های 1 تا n نیست، احتمال پیدا کردن توپ در جعبه z برابر است با

$$P = \frac{(1-r)P(A_z)}{r(1-P(A_z)-P(A_m)) + (1-r)(P(A_z)+P(A_m))}$$

و باز اگر همه جعبه‌ها را هم احتمال بگیریم، یعنی $\frac{1}{n+2}$ $P(A_z) = P(A_m) = P(A_i)$ خواهیم داشت

$$P = \frac{1-r}{nr+2(1-r)}.$$

برای مقایسه این نتیجه با نتیجه بخش قبل فرض می‌کنیم تعداد همه جعبه‌های در اختیار ما 10 تاست (یعنی $n = 9$) ولی تعداد کل جعبه‌ها درواقع 11 تاست و احتمال وجود توپ در این یازده جعبه یکسان است ($\frac{1}{11}$). احتمال خطای آزمایش را هم مانند قبل 0.1 می‌گیریم. در این صورت احتمال یافتن توپ در جعبه z برابر خواهد بود با $P = \frac{1}{3}$. به عبارت دیگر آن حالتی را که با روش حذف ناممکن‌ها حقیقت محض می‌دانیم فقط کمی بیش از 30 درصد احتمال دارد که حقیقت باشد!

سخن پایانی

دنیای واقعی بر خلاف دنیای قصه‌ها پر از عدم قطعیت، خطا و بی‌دقتی است. در چنین دنیایی دست یافتن به حقیقت به سادگی قصه‌ها نیست. بنابراین برای پرهیز از نتیجه‌گیری نادرست یا پیش‌بینی نادرست بهتر است تا حد امکان نگاه همه‌جانبه به پدیده‌ها داشته باشیم و امکان بروز خطا در مشاهده‌ها و آزمایش‌ها را نیز نادیده نگیریم.