
امید ریاضی

حسین نادری

دانش آموخته‌ی سمپاد
دانشجوی علوم کامپیوتر شریف

hnadri268.blog.ir
hnaderi268@gmail.com

۲۱ اسفند ۱۳۹۴

فهرست مطالب

۳	۱	شرط بندی روی سکه ها
۶	۲	تعریف های معادل برای امید ریاضی
۸	۱.۲	زمان متوسط انتظار قبل از خرابی برنامه
۹	۲.۲	بچه ی دختر
۹	۳	رفتار خطی امید ریاضی
۱۰	۱.۳	امید ریاضی عدد برآمده از دو تاس
۱۰	۲.۳	مساله تحویل کلاه
۱۲	۳.۳	پیش غذای چینی
۱۲	۴	بازی حدس عدد
۱۲	۱.۴	تحلیل بازی
۱۳	۵	تمبر جمع کن
۱۴	۱.۵	راه حل با کمک ویژگی خطی امید ریاضی
۱۵	۶	امید ریاضی ضرب چند متغیر تصادفی
۱۶	۱.۶	حاصلضرب دو تاس مستقل
۱۶	۲.۶	حاصلضرب دو تاس وابسته به هم
۱۷	۳.۶	چند نتیجه
۱۸	۱.۳.۶	پارادوکس RISC
۱۹	۲.۳.۶	برداشت مبتنی بر احتمال
۱۹	۳.۳.۶	مثال ساده تر
۲۰	۷	قصیه کلی امید ریاضی

امید ریاضی

امید ریاضی یک مقدار یا متغیر تصادفی، یک عدد است که اطلاعات زیادی درباره‌ی رفتار آن متغیر می‌دهد. اگر کمی نادقیق بخواهیم بیان کنیم، امید ریاضی میانگین ارزش هاست؛ به شرطی که وزن هر ارزش با احتمال رخ دادن آن سنجیده می‌شود. امید ریاضی متغیر R در فضای نمونه S به صورت زیر تعریف می‌شود:

$$Ex(R) = \sum_{\omega \in S} R(\omega) Pr(\omega)$$

برای آن که مفهوم امید ریاضی را کمی شهودی تر درک کنید، فرض کنید S مجموعه ای از دانش آموزان یک کلاس است و ما یک دانش آموز را با توزیع یکسان احتمال انتخاب می‌کنیم. R را نمره ی امتحان دانش آموز انتخاب شده در نظر می‌گیریم. حال $Ex(R)$ همان میانگین نمرات کلاس - اولین چیزی که هرکس بعد از پایان امتحان می‌خواهد بداند - است! به همین طریق امید ریاضی هم اولین چیزی است که یک نفر می‌خواهد در مورد یک متغیر تصادفی بداند. با یک مثال ادامه می‌دهیم. R را عدد برآمده از یک تاس منصف در نظر می‌گیریم. امید ریاضی R برابر است با:

$$\begin{aligned} Ex(R) &= \sum_{k=1}^6 k \left(\frac{1}{6}\right) \\ &= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} \\ &= 3\frac{1}{2} \end{aligned}$$

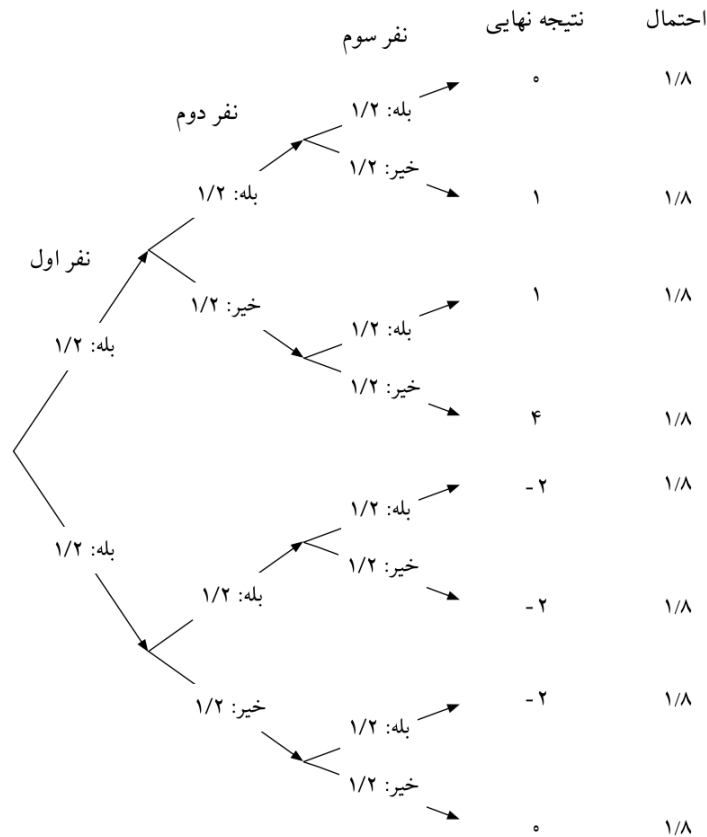
این محاسبه نشان می‌دهد عبارت "expected value" ^۱ می‌تواند همراه کننده باشد. چون متغیر تصادفی ممکن است هرگز نتواند برابر مقدار امید ریاضی شود. در حالی که عبارت فارسی امید ریاضی چندان مشکل ساز نیست!

۱ شرط بندی روی سکه ها

در یک بازی سه نفره، هر بازیکن ۲ سکه روی میز می‌گذارد و به طور مخفیانه روی یک کاغذ «شیر» یا «خط» می‌نویسد. سپس یکی از آن‌ها یک سکه منصف را به روی میز می‌اندازد. ۶ سکه ی روی میز به نسبت مساوی میان کسانی که روی کاغذشان درست پیش بینی کرده بودند، می‌رسد. اگر همه اشتباه حدس زده بودند، سکه های هرکس به خودش باز می‌گردد. پس از انجام چندین بازی، نفر اول تعداد زیادی سکه از دست می‌دهد. آنقدر زیاد که دیگر نمی‌توان گفت به خاطر شانس بدش بوده است. به نظر شما چه اتفاقی دارد می‌افتد؟

شکل ۱ نمودار درختی حالت های مساله با این فرض است که هر نفر به احتمال $\frac{1}{4}$ درست و مستقل از دیگران حدس می‌زند، است.

^۱ امید ریاضی به انگلیسی



شکل ۱: درخت تصمیم با فرضیات اولیه

ستون نتیجه نهایی برای نفر اول با احتساب دو سکه ای که برای شرکت در بازی پرداخت کرده، محاسبه شده است. امید ریاضی سود یا ضرر نفر اول^۲ را با توجه به نمودار حساب می کنیم:

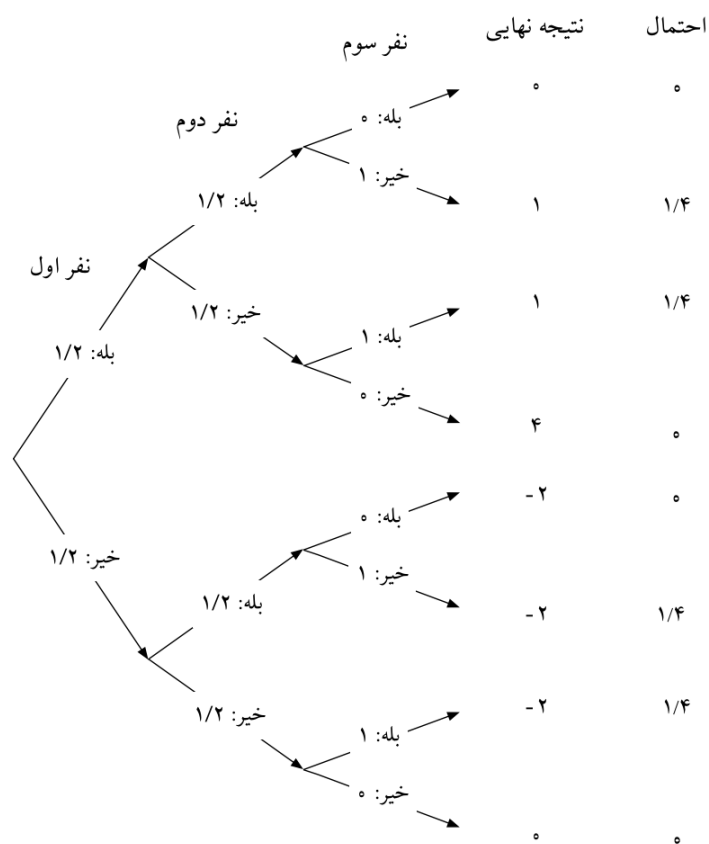
$$Ex(P) = 0 \times \frac{1}{8} + 1 \times \frac{1}{8} + 1 \times \frac{1}{8} + 4 \times \frac{1}{8} + (-2) \times \frac{1}{8} + (-2) \times \frac{1}{8} + 0 \times \frac{1}{8} = 0$$

پس بازی کاملاً منصفانه است و نفر اول پس از تعدادی بازی نه باید چیزی از دست بدهد و نه چیزی به دست آورد.

علت این مغایرت با آن چه در سوال مطرح شد، تبانی نفر دوم و سوم است. این دو نفر همیشه حدس های مخالف هم می زنند. و فرض ما مبنی بر این که هرکس مستقل از دیگران حدس می زند، نادرست

^۲ P حرف اول Payoff به معنی نتیجه نهایی است و برای حرف نشان دهنده ی متغیر تصادفی انتخاب شده است.

است. نفر اول به هیچ وجه نمی تواند همه ی سکه های روی میز را ببرد، چون حتی اگر ببرد، باید سکه های روی میز را با نفر دیگر که او نیز درست حدس زده تقسیم کند. به این ترتیب امید ریاضی او کم می شود و نمودار زیر به دست می آید:



شکل ۲: نمودار با در نظر گرفتن تبانی نفر دوم و سوم

از روی نمودار $Ex(P)$ را دوباره حساب می کنیم:

$$\begin{aligned}
 Ex(P) &= 0 \times 0 + 1 \times \frac{1}{4} + 1 \times \frac{1}{4} + 4 \times 0 + (-2) \times \frac{1}{4} + (-2) \times \frac{1}{4} + 0 \times 0 \\
 &= -\frac{1}{2}
 \end{aligned}$$

پس او به طور میانگین در هر بازی نیمی از یک سکه را از دست می دهد!

موقعیت های مشابه برای حقه های زیرکانه در بسیاری از بازی های شرط بندی وجود دارد. این مثال مربوط به Herman Chernoff^۳ می شود. قرعه کشی های دولتی بدترین نوع قمار بازی هستند؛ چون دولت تنها قسمتی از پول وارد بازی شده را به برندگان پرداخت می کند. با این حال Chernoff راهی برای بردن در قرعه کشی های دولتی نیز پیدا کرد. قوانین یک بازی مرسوم به این ترتیب است:

- هر بازیکن باید ۱ دلار بپردازد و ۴ عدد از ۱ تا ۳۶ انتخاب کند.
- دولت نیز ۴ عدد کاملاً تصادفی از ۱ تا ۳۶ انتخاب می کند.
- دولت نیمی از پول جمع شده را بین افرادی که درست حدس زده اند، تقسیم می کند و نیمی دیگر را صرف یک پروژه ساخت و ساز می کند.

این قرعه کشی وضعیت بسیار مشابهی با بازی شرط بندی روی سکه ها دارد، جز این که تعداد بازیکن ها و انتخاب ها بسیار بیشتر اند. Chernoff کشف کرد که مجموعه ی اعدادی که سهم بزرگی از جمعیت انتخاب میکنند، بسیار کوچک است^۴. این واقعیت مانند آن است که این افراد با هم همکاری می کنند که ببازند! حتی اگر یکی از آن ها درست حدس بزند، آن ها مجبور می شوند پول را میان عده ی زیادی که درست حدس زده اند تقسیم کنند. Chernoff با انتخاب تصادفی اعداد به احتمال کمی یکی از این دنباله های دوست داشتنی مردم را انتخاب می کند. پس اگر او ببرد، به احتمال بالایی کل پول نصیب او می شود! با تحلیل داده های واقعی قرعه کشی های دولتی او حساب کرد که می تواند به صورت میانگین به ازای هر دلار از این طریق ۷ سنت به دست آورد! کسانی که آشنایی با این بازی ها دارند، می دانند ۷ سنت به ازای یک دلار، رقم چندان بدی نیست.

۲ تعریف های معادل برای امید ریاضی

چندین روش برای نوشتن تعریف امید ریاضی وجود دارد. گاهی اوقات استفاده از یکی از صورت ها باعث می شود محاسبه بسیار آسان تر شود. یک روش به این صورت است:

قضیه ۱

$$Ex(R) = \sum_{x \in \text{range}(R)} x \cdot Pr(R = x)$$

^۳ استاد سابق آمار MIT

^۴ ظاهراً مردم شبیه به هم فکر می کنند یا به عبارتی حتی اگر افراد اعدادشان را تصادفی انتخاب کنند؛ انسان ها در ساختن اعداد تصادفی بسیار ضعیف عمل می کنند.

اثبات. سمت چپ را به سمت راست می‌رسانیم. $[R = x]$ را پیشامدی تعریف می‌کنیم که در آن $R = x$.

$$\begin{aligned} \text{Ex}(R) &= \sum_{\omega \in S} R(\omega) Pr(\omega) \\ &= \sum_{x \in \text{range}(R)} \sum_{\omega \in [R=x]} R(\omega) Pr(\omega) \\ &= \sum_{x \in \text{range}(R)} \sum_{\omega \in [R=x]} x Pr(\omega) \\ &= \sum_{x \in \text{range}(R)} \left(x \cdot \sum_{\omega \in [R=x]} Pr(\omega) \right) \\ &= \sum_{x \in \text{range}(R)} x \cdot Pr(R = x) \end{aligned}$$

در خط دوم یک جمع بندی را به دو جمع بندی شکستیم. جمع بیرونی تمامی مقدارهای ممکن x که متغیر تصادفی می‌تواند رخ دهد را در بر می‌گیرد و جمع درونی تمامی برآیندهایی که آن مقدار را دارند. بنابراین جمع ما هنوز کل فضای نمونه را می‌پوشاند. در خط آخر، از تعریف احتمال پیشامد $[R = x]$ استفاده کرده‌ایم.

◇

نتیجه ۲ اگر R یک متغیر تصادفی با مقدار طبیعی باشد، داریم:

$$\text{Ex}(R) = \sum_{i=0}^{\infty} i \cdot Pr(R = i)$$

اگر متغیر تصادفی فقط مقادیر طبیعی داشته باشد، راه دیگری نیز برای نوشتن امید ریاضی وجود دارد:

قضیه ۳ اگر R یک متغیر تصادفی با مقدار طبیعی باشد، داریم:

$$\text{Ex}(R) = \sum_{i=0}^{\infty} Pr(R > i)$$

اثبات. مجموع زیر را در نظر بگیرید:

$$\begin{aligned} &Pr(R = 1) + Pr(R = 2) + Pr(R = 3) + Pr(R = 4) + \dots \\ &+ Pr(R = 2) + Pr(R = 3) + Pr(R = 4) + \dots \\ &+ Pr(R = 3) + Pr(R = 4) + \dots \\ &+ Pr(R = 4) + \dots \end{aligned}$$

جمع ستون i ام برابر $i \cdot Pr(R = i)$ است. بنابراین جمع کل برابر است با:

$$\sum_{i=0}^{\infty} i \cdot Pr(R = i) = Ex(R)$$

از طرف دیگر سطر i ام برابر $\sum Pr(R > i)$ است. بنابراین جمع کل برابر است با:

$$\sum_{i=0}^{\infty} i \cdot Pr(R = i)$$

◇

۱.۲ زمان متوسط انتظار قبل از خرابی برنامه

مساله ای که الان مطرح می کنیم با یکی از تعریف های معادل امید ریاضی به آسانی حل می شود. یک برنامه کامپیوتری به احتمال p در پایان هر ساعت از کار می افتد. امید ریاضی زمان از کار افتادن برنامه چیست؟

اگر R تعداد ساعت هایی باشد که طول می کشد تا برنامه از کار بیفتد، جواب مساله $Ex(R)$ است. چون متغیر تصادفی طبیعی است، از رابطه ی زیر استفاده می کنیم:

$$Ex(R) = \sum_{i=0}^{\infty} Pr(R > i)$$

$R > i$ در صورتی رخ می دهد که برنامه بعد از i فرصت برای از کار افتادن خراب نشود، که احتمالش برابر $(1 - p)^i$ است. با جای گذاری در رابطه داریم:

$$\begin{aligned} Ex(R) &= \sum_{i=0}^{\infty} (1 - p)^i \\ &= \frac{1}{1 - (1 - p)} \\ &= \frac{1}{p} \end{aligned}$$

در خط دوم از حد مجموع استفاده کرده ایم. بنابراین اگر احتمال از کار افتادن برنامه در پایان هر ساعت ۱٪ باشد؛ زمان مورد انتظار برای از کار افتادن دستگاه برابر ۱۰۰ ساعت است.

۲.۲ بچه ی دختر

زوجی می خواهند که یک بچه ی دختر داشته باشند. به احتمال $\frac{1}{4}$ بچه ی آن ها مستقل از جنسیت بقیه ی بچه هایشان دختر می شود. اگر این زوج اصرار بر داشتن بچه ی دختر داشته باشند، انتظار چند پسر بچه را باید داشته باشند؟

این مساله حالتی از مساله ی قبل است. سوال «چند ساعت قبل از خرابی؟» با سوال «چند بچه تا بچه ی دختر؟» یکسان است. این جا $p = \frac{1}{4}$ است. با تحلیل این احتمال به جواب ۲ بچه می رسیم که چون آخرین بچه دختر است، آن ها باید انتظار یک بچه ی پسر نیز داشته باشند.

۳ رفتار خطی امید ریاضی

امید ریاضی از یک قانون شگفت انگیز خطی بودن پیروی می کند. این قانون می گوید امید ریاضی یک جمع برابر جمع امید ریاضی هاست.

قضیه ۴ (رفتار خطی امید ریاضی) برای هر جفت متغیر تصادفی R_1 و R_2 داریم:

$$\text{Ex}(R_1 + R_2) = \text{Ex}(R_1) + \text{Ex}(R_2)$$

اثبات. اگر S فضای نمونه باشد.

$$\begin{aligned}\text{Ex}(R_1 + R_2) &= \sum_{\omega \in S} (R_1(\omega) + R_2(\omega)) \cdot Pr(\omega) \\ &= \sum_{\omega \in S} R_1(\omega) Pr(\omega) + \sum_{\omega \in S} R_2(\omega) Pr(\omega) \\ &= \text{Ex}(R_1) + \text{Ex}(R_2)\end{aligned}$$

◇

ویژگی خطی بودن امید ریاضی را به هر مجموعه ی متناهی از متغیر های تصادفی تعمیم می دهیم.

نتیجه ۵ برای هر متغیر تصادفی $R_1, R_2, R_3, \dots, R_k$ داریم:

$$\text{Ex}(R_1 + R_2 + \dots + R_k) = \text{Ex}(R_1) + \text{Ex}(R_2) + \dots + \text{Ex}(R_k)$$

نتیجه ی خطی بودن امید ریاضی بسیار شگفت آور است، زیرا برخلاف بسیاری از قضیه های دیگر در احتمال، نیاز نیست متغیر های تصادفی از یک دیگر مستقل باشند و می توان بدون در نظر گرفتن مستقل بودن یا نبودن متغیر های تصادفی از آن استفاده کرد.

۱.۳ امید ریاضی عدد برآمده از دو تاس

امید ریاضی جمع عدد های برآمده از دو تاس منصف چند است؟
متغیر تصادفی R_1 برابر عدد تاس اول و R_2 عدد تاس دوم است. در ابتدای این مقاله گفتیم امید ریاضی یک تاس $\frac{1}{4} \times 3$ است. امید ریاضی جمع دو مقدار را با کمک گرفتن از ویژگی خطی بودن امید ریاضی حساب می کنیم.

$$\begin{aligned} \text{Ex}(R_1 + R_2) &= \text{Ex}(R_1) + \text{Ex}(R_2) \\ &= 3 \frac{1}{4} + 3 \frac{1}{4} \\ &= 7 \end{aligned}$$

توجه داشته باشید که فرض نکردیم دو تاس از یک دیگر مستقل اند. امید ریاضی جمع دو تاس ۷ است، حتی اگر آن ها به هم چسبیده باشند! در نتیجه از چسباندن دو تاس به هم، یک تاس منصف به دست می آید. حساب کردن این امید ریاضی از طریق رسم نمودار درختی بسیار پیچیده می شود، چون ۳۶ حالت وجود دارد و حتی اگر فرض نمی کردیم که تاس ها مستقل از یک دیگراند، حساب کردنش دیگر یک کابوس می شد!

۲.۳ مساله تحویل کلاه

در یک مهمانی شام n مرد شرکت کرده اند. هر کدام یک کلاه با خود آورده است. کلاه ها در طول شام قاطی می شوند و هر کس پس از شام یک کلاه تصادفی دریافت می کند. هر نفر به احتمال $\frac{1}{n}$ کلاه خود را دریافت می کند. امید ریاضی تعداد افرادی که کلاه خود را دریافت می کنند، چند است؟ بدون کمک گرفتن از ویژگی خطی امید ریاضی، پاسخ دادن به این سوال بسیار دشوار است. احتمالاً تلاش اول حل مساله این طور باشد که متغیر تصادفی R برابر تعداد افرادی است که کلاه های خود را دریافت کرده اند. می خواهیم $\text{Ex}(R)$ را حساب کنیم. از تعریف امید ریاضی داریم:

$$\text{Ex}(R) = \sum_{i=0}^{\infty} i \cdot Pr(R = i)$$

اکنون با مشکل مواجه شدیم، چون حساب کردن $Pr(R = i)$ خودش کار سختی است و بعد از آن ما نیاز داریم که مقدار حساب شده را در جمع بندی جای گذاری کنیم. از این گذشته نیاز داریم که احتمال هر جایگشتی از کلاه ها را ثابت در نظر بگیریم. برای مثال فرض کنیم احتمال وقوع همه ی جایگشت های کلاه ها با یکدیگر برابر اند.

در تلاش دوم سعی می کنیم از ویژگی خطی بودن امید ریاضی استفاده کنیم. R همان متغیر تصادفی ای است که در تلاش اول به کار بردیم. ترفند حل سوال این است که R را به صورت جمع چند متغیر شاخص می نویسیم. $R_i = 1$ را پیشامدی تعریف می کنیم که در آن مرد i ام کلاه خودش را پس می

^۵ به این نوع متغیر های تصادفی، شاخص می گویند.

گیرد و $R_i = 0$ پیشامدی که کلاه فرد دیگری را پس می‌گیرد. تعداد افرادی که کلاه خودشان را پس می‌گیرند برابر مجموع زیر است:

$$R = R_1 + R_2 + R_3 + \dots + R_n$$

این متغیرهای شاخص مستقل از یک دیگر نیستند. برای مثال اگر $n - 1$ مرد کلاه خودشان را پس بگیرند، نفر آخر چاره‌ای جز گرفتن کلاه خود ندارد. اما از آن جایی که قرار است از ویژگی خطی امید ریاضی استفاده کنیم، لازم نیست در مورد مستقل بودن یا نبودن نگران باشیم! برای حساب کردن امید ریاضی دو طرف معادله بالا از ویژگی خطی استفاده می‌کنیم:

$$\begin{aligned} \text{Ex}(R) &= \text{Ex}(R_1 + R_2 + \dots + R_n) \\ &= \text{Ex}(R_1) + \text{Ex}(R_2) + \dots + \text{Ex}(R_n) \end{aligned}$$

تنها چیزی که باقی مانده این است که امید ریاضی هر یک از متغیرهای کمکی را حساب کنیم. از یک واقعیت ابتدایی که ارزش به خاطر سپردن دارد استفاده می‌کنیم:

واقعیت ۶ امید ریاضی یک شاخص برابر احتمالی است که شاخص ۱ باشد. به نماد گذاری ریاضی:

$$\text{Ex}(I) = \text{Pr}(I = 1)$$

اثبات.

$$\begin{aligned} \text{Ex}(I) &= 1 \cdot \text{Pr}(I = 1) + 0 \cdot \text{Pr}(I = 0) \\ &= \text{Pr}(I = 1) \end{aligned}$$

◇

پس ما فقط نیاز داریم که $\text{Pr}(R_i = 1)$ را حساب کنیم که همان احتمال این است که مرد i ام کلاه خودش را پس بگیرد. چون احتمال هر مرد برای گرفتن کلاه خودش $\frac{1}{n}$ است. داریم:

$$\begin{aligned} \text{Ex}(R) &= \text{Ex}(R_1 + R_2 + \dots + R_n) \\ &= \text{Pr}(R_1 = 1) + \text{Pr}(R_2 = 1) + \dots + \text{Pr}(R_n = 1) \\ &= n \cdot \frac{1}{n} = 1 \end{aligned}$$

پس ما باید به طور متوسط انتظار داشته باشیم یک نفر کلاه خودش را دریافت کند. توجه کنید که ما فرض نکردیم که احتمال همه‌ی جایگشت‌های کلاه‌ها یکسان است. فقط نیاز داشتیم بدانیم که هر نفر به احتمال $\frac{1}{n}$ کلاه خودش را پس می‌گیرد. به این دلیل راه حل ما بسیار عمومی است، به مثال بعد توجه کنید.

۳.۳ پیش‌غذای چینی

n نفر در یک رستوران چینی دور یک میز غذای دایره‌ای نشسته‌اند. روی میز n نوع پیش‌غذای متمایز چیده شده است. هر نفر شروع به خوردن پیش‌غذایی که جلوی‌اش است می‌کند. پس از مدتی کسی میز را می‌چرخاند، به طوری که پیش‌غذایی تصادفی جلوی هر نفر قرار می‌گیرد. امید ریاضی افرادی که پیش‌غذای قبلی‌شان دوباره جلوی‌شان قرار بگیرد، چند است؟

این سوال یک حالت خاص از مساله تحویل کلاه است. در مساله تحویل کلاه تنها فرض کردیم که هر نفر به احتمال $\frac{1}{n}$ کلاه خودش را پس می‌گیرد. فراتر از این چیزی در مورد این که جایگشت‌ها چگونه اند فرض نکردیم. چون پیش‌غذاها شیف‌ت دوری نسبت به جایگاه قبلی‌شان داشته‌اند، یا به همه پیش‌غذای قبلی‌شان دوباره رسیده است یا به هیچ‌کس نرسیده. با این حال تحلیل قبلی مان باز هم کاراست و امید ریاضی تعداد افرادی که پیش‌غذای خودشان به خودشان برمی‌گردد ۱ است.

۴ بازی حدس عدد

قواعد یک بازی دو نفره در ادامه آمده است که یک ویژگی عجیب امید ریاضی را معلوم می‌کند. ابتدا نفر اول یک توزیع احتمال روی اعداد طبیعی در نظر می‌گیرد. توزیع احتمال هر توزیعی می‌تواند باشد. برای مثال ممکن است یک توزیع یکسان روی $1, 2, \dots, 6$ باشد، مانند یک تاس منصف. یا یک توزیع دو جمله‌ای روی $1, 2, \dots, n$ یا حتی می‌توان به هر عدد احتمالی بزرگتر از صفر داد، به شرطی که جمع همه‌ی احتمال‌ها یک شود.

پس از این، نفر دوم عدد z را بر حسب توزیع احتمال نفر اول انتخاب می‌کند. سپس نفر اول نیز به همین منوال عدد y_1 را تصادفی بر حسب توزیع احتمال انتخاب می‌کند. اگر y_1 بزرگتر از z باشد، بازی تمام می‌شود. در غیر این صورت اگر $y_1 \geq z$ بود، بازی با انتخاب y_2, y_3, y_4, \dots تا جایی ادامه می‌یابد که یک بزرگتر از z انتخاب شود. امید ریاضی تعداد اعداد انتخاب شده چه مقدار است؟

نفر اول دست کم یک عدد انتخاب می‌کند، بنابراین امید ریاضی قطعاً بزرگتر از یک است. جوابی حدود ۲ و ۳ معقول به نظر می‌رسد. اما ممکن است برای فردی سوال شود که آیا امید ریاضی به نحوه‌ی توزیع وابسته است یا نه؟ در ادامه به درستی یا نادرستی این شهود می‌پردازیم.

۱.۴ تحلیل بازی

تعداد اعدادی که باید انتخاب شوند تا بازی تمام شود، مقداری طبیعی است. و همان طور که دیدیم رابطه‌ی مناسبی برای امید ریاضی متغیر تصادفی طبیعی وجود داشت:

$$Ex(\# \text{ اعداد انتخاب شده}) = \sum_{k=0}^{\infty} Pr(\# \text{ اعداد انتخاب شده} > k) \quad (1)$$

اگر نفر اول z و نفر دوم اعداد y_1, y_2, \dots, y_k را انتخاب کرده باشد. دو حالت ممکن است وجود داشته باشد:

- اگر بزرگترین عدد میان انتخاب‌ها یکتا باشد، احتمال این که عدد انتخابی نفر دوم باشد، $\frac{1}{k+1}$ است. پس به احتمال $\frac{1}{k+1}$ بازی ادامه می‌یابد.

• اگر چندین بزرگترین عدد وجود داشته باشد، بازی با احتمالی بزرگتر از $\frac{1}{k+1}$ ادامه می یابد و باید نفر اول عدد انتخاب کند.

در هر دو حالت بازی با احتمال بیش از $\frac{1}{k+1}$ ادامه می یابد، و نفر اول نیاز دارد بیش از k عدد انتخاب کند. به عبارت دیگر:

$$Pr(\# \text{اعداد انتخاب شده} > k) \geq \frac{1}{k+1} \quad (2)$$

می توان نتیجه گرفت برای این که تعداد انتخاب های نفر اول کم شود، باید احتمال های اعداد بسیار کوچک باشند. برای مثال توزیع برابر روی مجموعه ی $\{1, 2, \dots, 10100\}$. در این حالت به ازای k های بزرگ، احتمال این که نفر اول بیش از k انتخاب برای پایان رساندن بازی داشته باشد بسیار نزدیک به $\frac{1}{k+1}$ است. برای مثال احتمال این که بیش از ۹۹ عدد انتخاب شود، تقریباً ۱٪ است. شهوداً به نظر می رسد عدد متوسطی برای تعداد اعداد انتخابی وجود داشته باشد. متأسفانه بر خلاف شهود، اثبات آسانی وجود دارد که امید ریاضی اعدادی که باید انتخاب شوند مستقل از توزیع بی نهایت است. ۲ را در ۱ جای گذاری می کنیم.

$$\begin{aligned} Ex(\# \text{اعداد انتخاب شده}) &\geq \sum_{k=0}^{\infty} \frac{1}{k+1} \\ &= \infty \end{aligned}$$

این پدیده می تواند آشفتگی بسیاری ایجاد کند! برای مثال فرض کنید که در یک شبکه هر بسته داده با احتمال $\frac{1}{k}$ بیشتر یا مساوی k مرحله به تاخیر می افتاد. ممکن است به نظر خوب برسد؛ به احتمال ۱٪ بسته ۱۰۰ مرحله یا بیشتر تاخیر دارد. ولی امید ریاضی زمان تاخیر برای بسته در حقیقت بی نهایت است!

نکته ای ضروری که باید به آن اشاره کرد؛ این است که هر متغیر تصادفی لزوماً یک امید ریاضی خوش تعریف ندارد. این ایده ممکن است در ابتدا آزاردهنده باشد ولی به خاطر بسپارید که امید ریاضی فقط یک میانگین وزن دار است. و مجموعه های بسیاری وجود دارند که میانگین به صورت رسمی ندارند، مانند:

$$\{1, -2, 3, -4, 5, -6, \dots\}$$

اکیدا باید در همه ی قضیه های مرتبط با امید ریاضی عبارت «...همه ی امکان های امید ریاضی وجود دارد.» مطرح شود. اما ما از بیان این فرض را به طور ضمنی صرف نظر می کنیم. خوش بختانه معمولاً متغیرهایی که امید ریاضی ندارند در تمارین و مسائل معمولی نیستند.

۵ تمبر جمع کن

در یک رستوران پس از خرید هر غذای بچه، یک ماشین اسباب بازی به مشتری هدیه داده می شود. n رنگ ماشین اسباب بازی به رنگ های آبی، سبز، قرمز، خاکستری، ... وجود دارند. رنگ ماشینی که پس از خرید به مشتری تحویل داده می شود، کاملاً تصادفی و مستقل از خرید های قبلی او است. یک

مرد هر روز از رستوران غذای بچه می خورد. امید ریاضی تعداد غذایی که باید این مرد بخرد تا از همه رنگ ها، حداقل یک ماشین داشته باشد، چقدر است؟^۶

سوال های ریاضی مشابه ممکن است ظاهر های متفاوت داشته باشند. برای مثال امید ریاضی تعداد افرادی که باید از آن ها آمارگیری شود؛ تا به ازای هر ماه دست کم یک از نفر متولدین آن ماه آمارگیری شده باشد، چقدر است. در ادامه این مقدار را هم حساب می کنیم. مساله کلی عموماً با نام تمبر جمع کن شناخته می شود.

۱.۵ راه حل با کمک ویژگی خطی امید ریاضی

ویژگی خطی امید ریاضی چیزی شبیه استقرای ریاضی و اصل لانه کبوتری است. یک ایده ساده که در انواع روش های مبتکرانه به کار می رود. اینجا نیز برای حل مساله تمبر جمع کن، یک روش زیرکانه برای بهره بردن از ویژگی خطی به کار می گیریم. فرض کنید ۵ رنگ وجود دارد و مرد خریدار دنباله ی زیر^۷ را دریافت می کند.

خاکستری نارنجی آبی نارنجی آبی قرمز سبز آبی سبز آبی
دنباله را به پنج قطعه افراز می کنیم.

$$\underbrace{\text{خاکستری آبی}}_{X_4} \quad \underbrace{\text{نارنجی آبی}}_{X_3} \quad \underbrace{\text{قرمز سبز}}_{X_2} \quad \underbrace{\text{سبز آبی}}_{X_1} \quad \underbrace{\text{آبی}}_{X_0}$$

هر قطعه جایی تمام می شود که یک ماشین جدید برای بار اول به دنباله اضافه می شود. و قطعه ی بعدی از عضو بعد از آن شروع می شود. به این ترتیب می توانیم مساله تمبر جمع کن را به چند قسمت ساده تر بشکنیم و پس از تحلیل هر قسمت از ویژگی خطی برای حساب کردن امید ریاضی کل استفاده کنیم.

به مساله کلی که قصد داریم n نوع ماشین جمع کنیم، باز می گردیم. X_k طول قطعه ی k ام است. تعداد کل ماشین هایی که باید گرفته شوند، برابر است با:

$$T = X_0 + X_1 + \dots + X_{n-1}$$

تمرکزمان را معطوف X_k ، طول قطعه ی k ام می کنیم. ابتدای قطعه ی k ام، حداکثر k رنگ مختلف وجود دارد و زمانی قطعه تمام می شود که یک ماشین با رنگ جدید افزوده شود. وقتی k رنگ ماشین انتخاب شده اند، یک ماشین تکراری به احتمال $\frac{k}{n}$ و یک ماشین جدید به احتمال $1 - \frac{k}{n}$ انتخاب می شود. بنابراین امید ریاضی ماشین هایی که باید قبل از یک ماشین جدید گرفته شود طبق مساله «زمان متوسط انتظار قبل از خرابی برنامه»^۸ $\frac{n}{n-k}$ به دست می آید. پس داریم:

$$\text{Ex}(X_k) = \frac{n}{n-k}$$

^۶ همانطور که مشاهده خواهید کرد، داستان این مساله مطابق نامش نیست. ولی به خاطر صورت داستانی اش که می توان آن را با جمع کردن کلکسیون تمبر بیان کرد، به نام تمبر جمع کن معروف شده است.
^۷ از سمت چپ بخوانید.

حال از ویژگی خطی استفاده می‌کنیم:

$$\begin{aligned} \text{Ex}(T) &= \text{Ex}(X_0 + X_1 + \dots + X_{n-1}) \\ &= \text{Ex}(X_0) + \text{Ex}(X_1) + \dots + \text{Ex}(X_{n-1}) \\ &= \frac{n}{n-0} + \frac{n}{n-1} + \dots + \frac{n}{1} \\ &= n \left(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{1} \right) \\ &= nH_n \end{aligned}$$

جمع بندی خط یکی مانده به آخر همان سری همساز^۸ است. می‌دانیم که مجموع سری همساز تا عضو n ام کمتر از $\ln n$ است.^۹

از این حل کلی برای پاسخ دادن به سوالات گسسته استفاده می‌کنیم. مانند: امید ریاضی تعداد بار تاس انداختن که همه ی اعداد ۱ تا ۶ ظاهر شوند، چقدر است؟

$$6H_6 = 14.7\dots$$

یا همان طور که قبلا اشاره کردیم، امید ریاضی تعداد افرادی که باید از آن‌ها آمارگیری کرد تا دست کم یک نفر به ازای هر ماه در میان نمونه آماری باشد؟

$$365H_{365} = 2364.6\dots$$

۶ امید ریاضی ضرب چند متغیر تصادفی

تا کنون به اندازه ی کافی به امید ریاضی مجموع پرداخته شده است. حال در مورد امید ریاضی حاصل ضرب چند متغیر تصادفی بحث می‌کنیم. اگر R_1 و R_2 مستقل باشند، امید ریاضی حاصل ضرب آن‌ها برابر حاصل ضرب امید ریاضی آن‌هاست.

قضیه ۷ برای متغیرهای تصادفی مستقل R_1 و R_2 داریم:

$$\text{Ex}(R_1 \cdot R_2) = \text{Ex}(R_1) \cdot \text{Ex}(R_2)$$

اثبات. سمت راست را به سمت چپ می‌رسانیم

$$\begin{aligned} \text{Ex}(R_1) \cdot \text{Ex}(R_2) &= \left(\sum_{x \in \text{range}(R_1)} x \cdot \text{Pr}(R_1 = x) \right) \cdot \left(\sum_{y \in \text{range}(R_2)} y \cdot \text{Pr}(R_2 = y) \right) \\ &= \sum_{x \in \text{range}(R_1)} \sum_{y \in \text{range}(R_2)} xy \text{Pr}(R_1 = x) \text{Pr}(R_2 = y) \\ &= \sum_{x \in \text{range}(R_1)} \sum_{y \in \text{range}(R_2)} xy \text{Pr}(R_1 = x \cap R_2 = y) \end{aligned}$$

^۸سری هارمونیک

^۹در سری دوم سوالات همین وبلاگ به این موضوع پرداخته شده است.

خط دوم حاصلضرب دو مجموع است. چون فرض کرده ایم R_1 و R_2 مستقل اند، مجاز به این کار بوده ایم. بر روی حاصلضرب xy دسته بندی می کنیم:

$$\begin{aligned}
 &= \sum_{z \in \text{range}(R_1 \cdot R_2)} \sum_{x,y: xy=z} xy Pr(R_1 = x \cap R_2 = y) \\
 &= \sum_{z \in \text{range}(R_1 \cdot R_2)} \left(z \sum_{x,y: xy=z} Pr(R_1 = x \cap R_2 = y) \right) \\
 &= \sum_{z \in \text{range}(R_1 \cdot R_2)} z \cdot Pr(R_1 \cdot R_2 = z) \\
 &= Ex(R_1) \cdot Ex(R_2)
 \end{aligned}$$

◇

۱.۶ حاصلضرب دو تاس مستقل

فرض کنید دو تاس منصف مستقل از هم داریم. امید ریاضی حاصلضرب دو عدد برآمده از آن ها چند است؟
 R_1 و R_2 عدد های برآمده از تاس ها اند. امید ریاضی حاصلضرب را به این صورت حساب می کنیم:

$$\begin{aligned}
 Ex(R_1 \cdot R_2) &= Ex(R_1) \cdot Ex(R_2) \\
 &= 3\frac{1}{2} \cdot 3\frac{1}{2} \\
 &= 12\frac{1}{4}
 \end{aligned}$$

خط اول از قضیه ۷ کمک گرفتیم و در خط بعد از امید ریاضی عدد برآمده از یک تاس که قبلا آن را حساب کرده بودیم، استفاده کردیم.

۲.۶ حاصلضرب دو تاس وابسته به هم

دو تاس غیر مستقل در نظر بگیرید؛ به طوری که نتیجه ی تاس دوم همواره مانند نتیجه ی تاس اول است. آیا این فرض نتیجه را تغییر می دهد؟ آیا فرض مستقل بودن برای قضیه ۷ لازم است؟ همانند قبل از R_1 و R_2 برای عدد های برآمده از تاس ها استفاده می کنیم. امید ریاضی را مستقیما

به طریق زیر حساب می کنیم:

$$\begin{aligned} \text{Ex}(R_1) \cdot \text{Ex}(R_2) &= \text{Ex}(R_1^2) \\ &= \sum_{i=1}^6 i^2 Pr(R_1 = i) \\ &= \frac{1^2}{6} + \frac{2^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{5^2}{6} + \frac{6^2}{6} \\ &= 15\frac{1}{6} \end{aligned}$$

خط اول از فرض یکسان بودن نتیجه ی دو تاس استفاده کردیم. سپس R_1^2 را طبق تعریف باز کردیم. مشاهده می کنیم که امید ریاضی تغییر کرده است. پس شرط مستقل بودن برای امید ریاضی ضرب ضروری است.

۳.۶ چند نتیجه

قضیه ۱ را می توان تعمیم داد.

نتیجه ۸ اگر $R_1, R_2, R_3 \dots R_n$ متغیرهای تصادفی مستقل باشند، داریم:

$$\text{Ex}(R_1 \cdot R_2 \cdot \dots \cdot R_n) = \text{Ex}(R_1) \cdot \text{Ex}(R_2) \cdot \dots \cdot \text{Ex}(R_n)$$

در اثبات این تعمیم از اصل استقرا و تعریف متغیرهای مستقل استفاده می شود. از بیان جزئیات صرف نظر می کنیم.

تا کنون درباره ی امید ریاضی حاصل جمع و حاصلضرب می دانیم. اما توصیف کردن امید ریاضی وارون چندان کار راحتی نیست. یک تلاش بی ثمر از این دست را در ادامه می آوریم.

نتیجه نادرست ۹ اگر R متغیر تصادفی باشد، داریم:

$$\text{Ex}\left(\frac{1}{R}\right) = \frac{1}{\text{Ex}(R)}$$

مثال نقض می زنیم. اگر R متغیر تصادفی ای باشد که به احتمال $\frac{1}{2}$ یک و به احتمال $\frac{1}{2}$ دو باشد، داریم:

$$\begin{aligned} \frac{1}{\text{Ex}(R)} &= \frac{1}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2} \\ &= \frac{2}{3} \\ \text{Ex}\left(\frac{1}{R}\right) &= \frac{1}{1} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \\ &= \frac{3}{4} \end{aligned}$$

دو مقدار به دست آمده برابر نیستند، پس نتیجه غلط است. اما در ادامه یک نتیجه غلط دیگر می آوریم که درستی آن را اثبات می کنیم!

نتیجه نادرست ۱۰ اگر $Ex(R/T) > 1$ آن گاه $Ex(R) > Ex(T)$.

اثبات. از سمت اگر شروع می کنیم، هر دو طرف را در $Ex(T)$ ضرب می کنیم، سپس از قضیه ۱ استفاده می کنیم:

$$\begin{aligned} Ex(R/T) &> 1 \\ Ex(R/T) \cdot Ex(T) &> Ex(T) \\ Ex(R) &> Ex(T) \end{aligned}$$

◇

این اثبات غلط، چند اشکال دارد! گام اول به شرطی درست است که $Ex(T) > 0$. مهم تر این که حق نداشتیم از قضیه ۱ در گام دوم استفاده کنیم. چون لزوماً R/T و T مستقل نیستند. متأسفانه با وجود این که نتیجه ۱۰ غلط است، عده ای از آن استفاده می کنند!

۱.۳.۶ پارادوکس RISC

داده های زیر از مقاله ای از چند استاد مشهور گرفته شده است. آن ها می خواستند نشان دهند برنامه ها روی پردازنده های RISC^{۱۰} عموماً سریع تر از پردازنده های CISC^{۱۱} اجرا می شوند. به این خاطر، آن ها جدولی مانند جدول زیر از مدت زمان اجرای برنامه ها تهیه کردند:

CISC/RISC	CISC	RISC	برنامه
۰/۸	۱۲۰	۱۵۰	جست و جوی رشته
۱/۵	۱۸۰	۱۲۰	آزمون بیتی
۰/۲	۳۰۰	۱۵۰	آکرمن
۰/۵	۱۴۰۰	۲۸۰۰	مرتب سازی بازگشتی
۱/۲			میانگین

نویسندگان از میانگین نرخ CISC بر RISC نتیجه گرفتند که در حالت کلی برنامه هایی که روی CISC اجرا می شوند، ۱/۲ برابر کند تر از RISC اند. ولی یک مشکل اساسی وجود دارد، آن هم این که اگر به جای ستون آخر، نرخ RISC بر CISC نوشته شود، به نتیجه ای دیگر می رسیم.

^{۱۰} کامپیوتر کم دستور: یک نوع معماری ساخت کامپیوتر یا ریزپردازنده می باشد. در این معماری به جای استفاده از دستورالعمل های خاص، که در سایر معماری ها مرسوم است، سعی می شود که از یک مجموعه دستورالعمل حداقلی و فوق العاده بهینه سازی شده استفاده شود.

^{۱۱} کامپیوتر با دستورهای پیچیده: نوعی معماری کامپیوتری است که در آن یک دستورالعمل می تواند چندین دستورالعمل سطح پایین دیگر را اجرا کند (مانند خواندن از حافظه، عملیات حسابی و ذخیره در حافظه).

RISC/CISC	CISC	RISC	برنامه
۱/۲۵	۱۲۰	۱۵۰	جست و جوی رشته
۰/۶۷	۱۸۰	۱۲۰	آزمون بیتی
۰/۵	۳۰۰	۱۵۰	آکرمن
۲	۱۴۰۰	۲۸۰۰	مرتب سازی بازگشتی
۱/۱			میانگین

با استدلالی کاملاً مشابه آن چه نویسندگان استفاده کردند، می توانیم نتیجه بگیریم که برنامه های RISC به اندازه ۱۰٪ طولانی تر از برنامه های CISC اند! به نظر شما چه اتفاقی افتاده است؟

۲.۳.۶ برداشت مبتنی بر احتمال

برای روشن شدن موضوع، رقابت بین RISC و CISC را با ابزارهای احتمال صورت بندی می کنیم. فضای نمونه، مجموعه ی برنامه های بررسی شده اند. متغیر تصادفی R طول برنامه های RISC و متغیر تصادفی C طول برنامه های CISC اند. می خواهیم میانگین مدت زمان اجرای برنامه های RISC یعنی $Ex(R)$ را با میانگین مدت زمان اجرای برنامه های CISC یعنی $Ex(C)$ مقایسه کنیم. برای مقایسه طول برنامه ها باید برای هر نمونه یک احتمال نسبت دهیم. ممکن است کسی بخواهد بر اساس نرخ استفاده ی برنامه های مشابه به احتمال ها وزن دهد. اما ما از روش نویسنده ها پیروی می کنیم. آن ها وزن های برابر به برنامه ها دادند، پس به طور ضمنی فرض کرده اند که برنامه های مشابه ارزش های یکسان دارند. همین کار را می کنیم و $Ex(R)$ و $Ex(C)$ را به این طریق به دست می آوریم:

$$Ex(R) = \frac{150}{4} + \frac{120}{4} + \frac{150}{4} + \frac{2800}{4}$$

$$= 805$$

$$Ex(C) = \frac{120}{4} + \frac{180}{4} + \frac{300}{4} + \frac{1400}{4}$$

$$= 500$$

امید ریاضی طول برنامه های RISC دقیقاً $1/61$ برابر طول برنامه CISC است. RISC حتی بدتر از دو پاسخ پیشنهادی قبل عمل می کند!

بر طبق مدل احتمالاتی ما، نویسنده ها C/R را برای هر نمونه حساب کردند و با میانگین گرفتن از آن ها نتیجه گرفتند $Ex(C/R) = 1/2$. تا این جا درست پیش رفتند. اما آن ها این گونه برداشت کردند که برنامه های CISC کندتر از برنامه های RISC اند. از این قرار، کلید نتیجه ی غلط مقاله متکی به نتیجه ۱۰ است^{۱۱}، که همه علت غلط بودن آن را می دانیم!

۳.۳.۶ مثال ساده تر

خاستگاه مساله در آن چه در زیر آمده است، روشن تر است. فرض کنید داده ها از این دست باشند:

^{۱۱} چون به جای تقسیم امید ریاضی طول برنامه ها، از امید ریاضی تقسیم متغیرهای تصادفی بر یک دیگر استفاده کرده بودیم.

B/A	A/B	پردازنده B	پردازنده A	معیار
۲	۱/۲	۲	۱	مساله ۱
۱/۲	۲	۱	۲	مساله ۲
۱/۲۵	۱/۲۵			میانگین

حال آمار و ارقام مربوط به پردازنده A و B متقارن می باشند. با این حال می توان از ستون سوم نتیجه گرفت پردازنده B، ۲۵% کند تر از A است و از ستون چهارم می توان نتیجه گرفت پردازنده A، ۲۵% کند تر از B است. جفت نتیجه گیری ها به وضوح غلط اند. مفهوم میانگین های نرخ ها گمراه کننده است. از این پس هرگاه دارید امید ریاضی یک کسر را حساب کنید، دو بار فکر کنید؛ تا از یک مقدار خام استفاده و برداشت اشتباه نکنید.

۷ قضیه کلی امید ریاضی

همان طور که احتمال شرطی وجود دارد، امید شرطی نیز وجود دارد. برای مثال، امید ریاضی عدد بر آمده از تاس به شرطی که عدد بر آمده زوج باشد را در نظر بگیرید. همان روش های محاسبه امید ریاضی معمولی برای امید ریاضی شرطی نیز صادق است. با این تفادیت که همه ی احتمال ها به احتمال های شرطی تبدیل می شوند. اگر R متغیر تصادفی و E یک پیشامد باشد، امید ریاضی به شرطی که پیشامد E روی دهد با $Ex(R | E)$ نمایش داده می شود و به این صورت تعریف می شود:

$$\begin{aligned} Ex(R | E) &= \sum_{\omega \in S} R(\omega) Pr(\omega | E) \\ &= \sum_{x \in \text{range}(R)} x \cdot Pr(R = x | E) \end{aligned}$$

به عنوان مثال، اگر R عدد بر آمده تاس و E رویدادی باشد که عدد بر آمده زوج باشد. امید $Ex(R|E)$ را حساب می کنیم.

$$\begin{aligned} Ex(R | E) &= \sum_{\omega \in \{1, \dots, 6\}} R(\omega) Pr(\omega | E) \\ &= 1 \times 0 + 2 \times \frac{1}{3} + 3 \times 0 + 4 \times \frac{1}{3} + 5 \times 0 + 6 \times \frac{1}{3} \\ &= 4 \end{aligned}$$

امید شرطی برای تفکیک محاسبه به چند حالت بسیار کارآمد است. تفکیک به وسیله ی قیاس کردن با قضیه کلی امید ریاضی به دست می آید.

قضیه ۱۱ (قضیه کلی امید ریاضی) E_1, E_2, \dots, E_n رخداد هایی باشند که فضای نمونه را افراز می کنند و احتمال های ناصفر دارند. اگر R متغیر تصادفی باشد، داریم:

$$Ex(R) = Ex(R | E_1) \cdot Pr(E_1) + \dots + Ex(R | E_n) \cdot Pr(E_n)$$

به عنوان مثال اگر R عدد برآمده از تاس و E پیشامدی باشد که عدد برآمده زوج و \bar{E} پیشامدی که عدد برآمده فرد باشد، باشند. امید ریاضی کلی برابر است با:

$$\underbrace{\text{Ex}(R)}_{\vee/2} = \underbrace{\text{Ex}(R | E)}_{=4} \cdot \underbrace{\text{Pr}(E)}_{=1/2} + \underbrace{\text{Ex}(R | \bar{E})}_{=?} \cdot \underbrace{\text{Pr}(\bar{E})}_{=1/2}$$

تنها کمیتی که قبلا مقدارش را نمی دانستیم $\text{Ex}(R | \bar{E})$ که امید ریاضی با شرط فرد بودن عدد برآمده است. با حل این معادله به $\text{Ex}(R | \bar{E}) = 3$ می رسیم. برای اثبات قضیه کلی امید ریاضی ابتدا لم زیر را اثبات می کنیم.

لم ۱۲ اگر R یک متغیر تصادفی، E یک پیشامد با احتمال بزرگتر از صفر و I_E متغیر شاخص برای E باشد. داریم:

$$\text{Ex}(R | E) = \frac{\text{Ex}(R \cdot I_E)}{\text{Pr}(E)} \quad (3)$$

اثبات. دقت داشته باشید که برای هر برآمد s در فضای نمونه :

$$\text{Pr}(\{s\} \cap E) = \begin{cases} 0 & I_E(s) = 0 \\ \text{Pr}(s) & I_E(s) = 1 \end{cases}$$

و هم چنین:

$$\text{Pr}(\{s\} \cap E) = I_E(s) \cdot \text{Pr}(s) \quad (4)$$

حال،

$$\begin{aligned} \text{Ex}(R | E) &= \sum_{s \in S} R(s) \cdot \text{Pr}(\{s\} | E) && (\text{طبق تعریف } \text{Ex}(\cdot | E)) \\ &= \sum_{s \in S} R(s) \cdot \frac{\text{Pr}(\{s\} \cap E)}{\text{Pr}(E)} && (\text{طبق تعریف } \text{Pr}(\cdot | E)) \\ &= \sum_{s \in S} R(s) \cdot \frac{I_E(s) \cdot \text{Pr}(s)}{\text{Pr}(E)} && ((4) \text{ طبق}) \\ &= \frac{\sum_{s \in S} (R(s) \cdot I_E(s)) \cdot \text{Pr}(s)}{\text{Pr}(E)} \\ &= \frac{\text{Ex}(R \cdot I_E)}{\text{Pr}(E)} && (\text{طبق تعریف } \text{Ex}(R \cdot I_E)) \end{aligned}$$

◇

اکنون قضیه کلی امید ریاضی را اثبات می‌کنیم.
اثبات. چون E_i ها فضای نمونه را افراز می‌کنند،

$$R = \sum_i R \cdot I_{E_i} \quad (5)$$

برای هر متغیر تصادفی R داریم:

$$\begin{aligned} \text{Ex}(R) &= \text{Ex}\left(\sum_i R \cdot I_{E_i}\right) && \text{((5) طبق)} \\ &= \sum_i \text{Ex}(R \cdot I_{E_i}) && \text{(ویژگی خطی امید ریاضی)} \\ &= \sum_i \text{Ex}(R | I_{E_i}) \cdot Pr(E_i) && \text{((3) طبق)} \end{aligned}$$

◇