احتمال، بسامد و انتظار معقول

آر. تی. کاکس [1]، مترجم: محمدرضا مشکانی [2]

یادداشت:

در سال‌هایی که هنوز نظریهٔ آمار بیزی با اقبال عام روبه‌رو نشده بود و حرف و حدیث فراوان دربارهٔ آن رایج بود، فرهیخته‌ای از تبار فیزیکدانان به نام آر. تی. کاکس نظریهٔ احتمال بیزی را به صورت اصل موضوعی اثبات کرد. سال‌ها بعد دانشمندی دیگر با شرح و بسط بیشتر کار کاکس آن را دست یافتنی‌تر ساخت.

براین باورم که آگاهی همکاران و دانشجویان فارسی زبان از این کوشش‌ها سودمند تواند بود، زیرا اطمینان خواهند یافت که نظریهٔ آمار بیزی بر پایهٔ ریاضی محکم استوار است. با این باور، ترجمهٔ این مقاله‌های ارزشمند به خوانندگان اندیشهٔ آماری تقدیم می‌شود. باشد که با تأمل و تعمق در این آثار در امور علمی خود از نتیجه‌های آنها بهره‌مند شوند.

محمدرضا مشکانی

مشخصات مقاله اصلی:

Cox, R. T. (1946). Probability, Frequency, and Reasonable Expectation. *American Journal of Physics*: **14(1)**, pp. 1-13.

۱ بسامد یا انتظار معقول به‌عنوان مفهوم بنیادی

مفهوم احتمال از بدو شروع نظریهٔ احتمال دو اندیشه را دربرداشت: اندیشهٔ بسامد در یک مجموعه و اندیشهٔ انتظار معقول. گزینش یکی یا دیگری از این دو به‌عنوان معنای بنیادین احتمال موجب تمایز دو مکتب فکری در این نظریه شده است[3].

اگر جعبه‌ای دو گوی سفید و یک گوی سیاه را دربر داشته باشد که به جز رنگ‌شان غیر قابل تمیزند، هر دو مکتب موافق‌اند که احتمال آنکه شخصی با چشم بسته در یک بار استخراج گویی سفید بیرون آورد $\frac{2}{3}$ و احتمال آنکه گویی سیاه بیرون بیاورد $\frac{1}{3}$ است. در نظریهٔ بسامدی، معنای اصلی این احتمال‌ها برحسب هماد[4] بیان می‌شود. هماد ممکن است شماری به‌طور نامحدود بزرگ از چنین جعبه‌های دارای همان محتوا باشد، یا ممکن است شماری به‌طور نامحدود بزرگ از استخراج از یک جعبه باشد که هر بار گوی بیرون آورده شده به جعبه برگردانده می‌شود. نکتهٔ مهم این است که فرض می‌شود شرایط آغازین مستعد تکرارهای نامحدودند، این تکرارها هماد را تشکیل می‌دهند. این گفته که احتمال گوی سفید $\frac{2}{3}$ است صرفاً بدین معنی است که شمار آزمایه‌هایی که گوی سفید را به‌عنوان نتیجهٔ استخراج به دست می‌دهند $\frac{2}{3}$ شمار آزمایه‌های کل هماد ست. طبق نظریهٔ بسامدی، این یک پیشگویی نظریهٔ احتمال نیست بلکه تعریف احتمال است. احتمال در آن نظریه سرشتی از هماد است و بدون هماد نمی‌توان گفت که وجود دارد.

باز هم بنابر هر دو مکتب، احتمال داشتن دو گوی سفید در دو استخراج پیاپی، وقتی که گوی اول بیرون آورده شده جایگذاری نمی‌شود، برابر با $\frac{1}{2} \times \frac{2}{3}$ یا $\frac{1}{3}$ است. طبق نظریهٔ بسامدی، این گفته

[1] دانشگاه جانز هاپکینز، بالتیمور ۱۸،مریلند

[2] دانشگاه شهید بهشتی، گروه آمار

[3] اگر تفاوت‌های جزئی را به شمار آوریم، به‌نظر می‌رسد شمار مکتب‌ها عددی بین دو و تعداد مؤلفان است و احتمالاً به عدداخیر نزدیک‌تر است. اما آشکارترین خط تقسیم همان است که ذکر شد.

[4] Ensemble

بدان معنی است که از هر یک از همادی از جعبه‌های بدواً دارای دو گوی سفید و یک گوی سیاه، دو گوی به طور پیاپی استخراج می‌شوند. در $\frac{2}{3}$ آزمایه‌ها، اول گویی سفید استخراج می‌شود و یک گوی سفید و یک گوی سیاه در جعبه باقی می‌مانند. سپس در $\frac{1}{2}$ آزمایه‌هایی که این نتیجه را به دست داده‌اند، گویی سفید در مرحلهٔ بعد بیرون آورده می‌شود، به‌گونه‌ای که $\frac{1}{3}$ کل شمار آزمایه‌ها در هر دو استخراج گوی‌های سفید را به دست می‌دهند. مثال‌های بالا این واقعیت کلی را نشان می‌دهند که وقتی احتمال با مفهوم بسامد در یک هماد تعریف می‌شود، مقدارهای احتمال در مثال‌های خاص به وسیلهٔ عملیات حسابی و در حالت‌های کلی قاعده‌های احتمال به کمک محاسبات جبر معمولی به دست آورده می‌شوند[۵] [۱].

احتمال، به عنوان فراهم آورندهٔ معیاری از انتظار معقول از رخداد یک پیشامد در آزمایه‌ای تک نیز به رسمیت شناخته می‌شود. این که احتمال بیرون آوردن گویی سفید $\frac{2}{3}$ و بیرون آوردن گوی سیاه $\frac{1}{3}$ است بدان معناست که بیرون آمدن گویی سفید نتیجهٔ محتمل‌تری از یک آزمایه در قیاس با بیرون آمدن گویی سیاه است و عددهای $\frac{2}{3}$ و $\frac{1}{3}$ برای مقایسهٔ درستنمایی‌های این دو نتیجه به کار می‌آیند. طبق دومین مکتب عمدهٔ احتمال، این معیار انتظار معقول، به جای بسامد در یک هماد، معنای بنیادی احتمال است.

اگر بتوان نشان داد که هر ملاکی از انتظار معقول نیز یک بسامد در نوعی هماد است و هر بسامد در یک هماد انتظاری معقول را می‌سنجد، آن‌گاه گزینش یکی یا دیگری به عنوان معنای بنیادی احتمال خیلی اهمیت نخواهد داشت. سعی نخواهم کرد در این باره بحث کنم که آیا بسامدهایی درهمادی وجود دارند که معیارهای انتظار معقول نیستند؟ برای هدف فعلی من کافی است نشان داده شود که این دو تفسیر همیشه یکسان نیستند. برای این کار اشاره به این نکته بسنده خواهد بود که احتمال‌هایی به معنای انتظارهای معقول وجود دارند که برای آنها هیچ همادی وجود ندارد واگر برای آنها همادی تصور شود، آشکارا چیزی بیش از نیرنگ ذهنی مناسب نیست. بدین ترتیب زمانی که احتمال وجود بیش از یک منظومهٔ سیارات در کیهان حساب می‌شود، حتی به عنوان یک نیرنگ به سختی قابل دفاع است که این احتمال به شماری از کیهان‌های دارای بیش از یک منظومهٔ سیارات در بین تعداد زیادی از کیهان‌هایی که همهٔ آنها به طریقی مشابه این کیهان‌اند و بنا به تعریف همه چیز را در بردارند، اشاره داشته باشد.

افزون برآن، چنان گذار تدریجی از حالت‌هایی که در آنها همادی قابل کشف وجود دارد به حالت‌هایی که در آنها چنین چیزی نیست، در بین است که نظریه‌ای که به تمایزی آشکار بین آنها نیازمند است مشکلاتی جدی پیش می‌نهد. مثالی چند این نکته را روشن خواهد ساخت. بگذارید احتمال آنکه شمار شیرهای ظاهر شده در تعدادی از پرتاب‌های یک سکهٔ سالم درون دو حدّ معینی قرار گیرد را در نظر بگیریم، و بگذارید کاری را که اغلب انجام می‌شود، یعنی احتمال آنکه یک ثابت فیزیکی درون دو حد معین واقع باشد را با احتمال شیرها مقایسه کنیم. این دو احتمال چیزی مشترک دارند، اما بین آنها تفاوتی وجود دارد. تفاوت در علت‌هایی است که هنگام بحث دربارهٔ امتیازهای به دست آمده از پرتاب یک سکه و مقدار یک ثابت فیزیکی، ما را وامی‌دارند به جای قطعیت‌ها با احتمال‌ها سر و کار داشته باشیم. هنگام بحث دربارهٔ امتیازهای حاصل از شمار معینی از پرتاب‌های یک سکه، ناچاریم احتمال را به کار گیریم زیرا امتیاز آمدن شیر از یک پرتاب به پرتاب دیگر تغییر خواهد کرد. از سوی دیگر، مقدار واقعی یک ثابت فیزیکی مقداری یکتاست. ناچاریم تنها به این دلیل که این مقدار بین دو حدّ معین واقع است از احتمال سخن بگوییم زیرا دانش ما ناکامل است.

گاهی، درست است که احتمال واقع شدن یک ثابت فیزیکی بین دو حدّ معین معادل با احتمالی دیگر است مبنی بر این که خطای موجود در متوسط شماری از اندازه‌گیری‌ها بین این دو حد واقع باشد. اگر منبع‌های خطای راژمند (سیستماتیک) در بین نباشند، می‌توانیم همادی از اندازه‌گیری‌ها را تصور کنیم و با اندازه‌گیری‌های به عمل آمده به مثابه نمونه‌ای تصادفی از این هماد رفتار کنیم. دراین صورت احتمال مورد بحث می‌تواند به روالی مشابه با روال مورد استفاده در رفتار با سکه به دست آورده شود. برای مثال، احتمال حدود معین برای مقدار واقعی معادل ژول را شاید بتوان به این طریق در نظر گرفت.

[۵] گزارشی از نظریهٔ بسامدی با تذکراتی دربارهٔ نظریه‌های دیگر توسط G. Bergman, (1941). American Journal of Physics. **9**, 263. ارائه شده است. خوانندگانی که در قیاس با دانش من دانش گسترده‌تری از فلسفه دارند بهتر خواهند توانست دیدگاه‌های وی را با دیدگاه‌های این مقاله مقایسه کنند.

وضعیت در خصوص وارون ثابت ریزساختار که در مکانیک کوانتومی ظاهر می‌شود تا اندازه‌ای متفاوت است. زیرا در اینجا، علاوه بر مقدارهای به دست آمده از اندازه‌گیری‌ها، شواهدی از نوع دیگر که در استدلالی که ادینگتون[٦] [۳] عرضه داشته است وجود دارند که می‌توان انتظار داشت این ثابت عددی صحیح برابر ۱۳۷ باشد. اگر باید تنها از اندازه‌گیری‌ها برآورد شود که احتمالی برابر وجود دارد که این ثابت درون یا بیرون حدود معین دربردارندهٔ ۱۳۷ واقع باشد، آن‌گاه استدلال ادینگتون احتمال آن را که درون این حدود باشد افزایش داده و در نتیجه احتمال آن را که بیرون آن واقع باشد، کاهش خواهد داد.

به عنوان مثال نهایی، می‌توانیم وضعیت یک ثابت صرفاً ریاضی را که وجودش ثابت شده اما مقدار آن تنها درون حدود معینی تعیین شده است در نظر بگیریم. مسئله‌ای از نظریهٔ اعداد که توسط اشخاص مختلف از جمله هاردی[۷] [٤] مورد بحث واقع شده، مثال خوبی را فراهم می‌سازد. این مسئله به برابری یک عدد صحیح با مجموع توان سوم عددهای صحیح کوچکتر از آن مربوط است. ثابت شده است که هر عدد صحیح برابر است با مجموع توان سوم ۹ عدد یا کمتر، و هر عدد صحیح بزرگتر از عددی معین برابر است با مجموع توان سوم ۸ عدد یا کمتر. انتظار می‌رود که همهٔ عددهای صحیح را که به قدر کافی بزرگ‌اند بتوان بامجموعی از تعداد باز هم کوچکتر از توان‌های سوم بیان کرد، و مسئله یافتن مینیمم شمار لازم برای همهٔ عددهای صحیح بزرگتر از مقداری معین است. ثابت شده است که اگر این مقدار به قدر کافی بزرگ اختیار شود، شمار مورد بحث ٤، ۵، ٦، ۷، یا ۸ است. برهان دقیق تا بدین‌جا پیش رفته است، اما شواهد حاصل از محاسبات، برخی از این شمارها را از دیگران کمتر محتمل می‌سازد. در نمونه‌های خیلی بزرگ - همهٔ ٤۰۰۰۰ عدد صحیح و ۲۰۰۰ عدد منتهی به یک میلیون - دیده می‌شود که به تدریج که به سوی عددهای بزرگتر پیشروی می‌شود، شمار عددهایی که توان سوم ۸ عدد را لازم دارند درهمان اوایل کاهش می‌یابد و عددهایی که به توان سوم ۷ عدد نیاز دارند قدری دورتر ناپدید می‌شوند، در حالی که اعدادی که به توان سوم ٦ عدد احتیاج دارند بیشتر رخ می‌دهند و با ندرت بیشتر دیده می‌شوند تا آنکه تنها به دو مورد در بین ۲۰۰۰ عدد صحیح بلافاصله زیر یک میلیون می‌رسیم. هاردی نتیجه می‌گیرد که شمار مینیمم برای عددهای صحیح به قدر کافی بزرگ تقریباً به یقین نه ۸، نه ۷ و احتمالا نه ٦ است. شمارهای ۵ و ٤ به عنوان شمارهای محتمل باقی می‌مانند و به نظر می‌رسد که وی با ٤ به عنوان شمار محتمل‌تر موافق است.

اکنون این چهار مثال را در نظر بگیریم: احتمال حدود معین برای (i) امتیاز حاصل از شماری از پرتاب‌های یک سکه، (ii) مقدار معادل ژول، (iii) مقدار وارون ثابت ریزساختار، و (iv) مقدار کوچکترین شمار از توان‌های سوم برای بیان عددهای صحیح بزرگ. احتمال زیاد می‌رود که مثال‌هایی دیگر را بتوان بین اینها جای داد به طوری که تفاوت‌های بین هر مثال و مثال‌های قبل و بعد آن بسیار اندک باشند. در این صورت سلسله‌ای مدرج از مثال‌های احتمال خواهیم داشت. در یک انتهای این سلسله، تفسیر احتمال برحسب بسامد معتبر خواهد بود، در انتهای دیگر چنین کاری ناممکن خواهد بود. زیرا یقیناً ناممکن است که دربارهٔ پراکنش آماری موارد تعیین عددی که در واقع هرگز تعیین نشده است و هنگامی که تعیینی از آن صورت گیرد، مقداری تک و منطقاً قطعی به دست خواهد داد، بحث کرد.

با وجود این، باید اعتراف کرد که نوعی استدلال مشترک برای همهٔ این مثال‌ها وجود دارد. قمار باز در مثال اول، فیزیکدان در مثال‌های دوم و سوم و ریاضیدان در مثال چهارم همه از فرایندهای استنباطی مشابه استفاده می‌کنند. در این خصوص ارزشمند است ملاحظه کنیم چقدر از نظریهٔ احتمال به بررسی روابط بین احتمال‌ها: بین احتمال آنکه پیشامدی رخ نخواهد داد و احتمال آنکه رخ خواهد داد، بین احتمال هر دو پیشامد و احتمال‌های جداگانهٔ آنها، بین این احتمال‌ها و احتمال آنکه دست کم یکی از این دو پیشامد رخ خواهد داد، می‌پردازد. در حالتی که احتمال‌ها را بتوان با بسامدهای موجود در یک هماد یکی پنداشت، همان‌گونه که پیشتر ذکر شد، این روابط به کمک جبر معمولی به آسانی به دست آورده می‌شوند. اما در استنباط ناظر به انتظار معقول، حتی وقتی هیچ همادی قابل کشف نیست،همین روابط یا دست کم روابط مشابه

[٦] A. S. Eddington, Relative Theory of Protons and Electrons (Macmillan, New York, and University Press, Cambridge, 1936).

[۷] G. H. Hardy,Some famous Problems of the Theory of Numbers and in Particular Waring's Problem (Clarendon Press, Oxford, 1920).

مطرح‌اند. بدین ترتیب، تحت هر تعریفی از احتمال، یا حتی بدون کوششی برای تعریف آن به طرزی دقیق، هنوز هم این توافق وجود خواهد داشت که پیشامدی که رخداد آن کمتر محتمل است، بیشتر محتمل است که رخ ندهد. رخداد توأم هر دو پیشامد محتمل‌تر از رخداد یکی از آن دو پیشامد که کمتر محتمل است، نخواهد بود و به طور کلی کمتر محتمل خواهد بود. اما رخداد دست کم یکی از آن پیشامدها از رخداد هر یک از آن دو پیشامد کمتر محتمل نیست و به طور کلی بیشتر محتمل است.

برای مثال، اگر مسئله میزان باورمندی یک فرضیهٔ معین مربوط به منشأ حیات روی زمین یا زبان انسان باشد، می‌توان علیه آن فرضیه‌ای که رخداد دو پیشامد را مسلّم انگاشته که هیچ یک از آن دو خیلی محتمل تلقی نمی‌شده به مثابه یک نکته منفی معتقد بود، اما این فرضیه اگر بتواند با مسلّم انگاشتن اینکه صرفاً یکی یا دیگری از این دو پیشامد رخ داده بوده، توجیه شود در کسب باورمندی توفیق خواهد یافت.

به بیان کلی، فرضیه ساده بر فرضیه پیچیده ترجیح دارد. اگر این ترجیح به جای قراردادی صرف برپایهٔ باوری معقول استوار باشد، به نظر می‌رسد اثبات‌اش این خواهد بود که دو یا چند فرضیه کمتر از یک فرضیه تک دربارهٔ یک درستنمایی، محتمل خواهند بود.

این مشکل نظریهٔ بسامدی احتمال را اکنون می‌توان چنین خلاصه کرد. عرصه‌ای از استنباط محتمل وجود دارد که بیرون از حیطهٔ آن نظریه است. اشتقاق قاعده‌های احتمال از مشخصه‌های هماد به کمک جبر عادی نمی‌تواند کاربست این قاعده‌ها را در این عرصه بیرونی توجیه کند. با این وجود، کاربست این قاعده‌ها در این عرصه همه جایی است و به نظر می‌رسد بخشی بنیادین از استدلال‌مان است. بدین سان نظریهٔ بسامدی از این لحاظ نارساست که نمی‌تواند آنچه را تصور می‌شود کاربست مشروع قاعده‌های خودش باشد، توجیه کند.

از دیدگاهی صرفاً عقلانی، گسترش این حیطه استنباط به بیرون از حیطهٔ نظریهٔ بسامدی برای نکتهٔ مورد پرسش بی‌ربط است. حتی اگر موارد معتبر استدلال در این عرصه کمیاب و کم اهمیت می‌بودند، هنوز هم منطقاً ضروری بود که به نارسایی نظریهٔ بسامدی باور داشته باشیم. اما به عنوان یک موضوع عملی،اگر این موارد اندک‌شمار یا پیش پا افتاده می‌بودند، احتمالاً باید قانع می‌شدیم که نادیده‌شان بگیریم. اما در حقیقت همان طور که سعی کرده‌ام با مثال‌های داده شده القاء کنم بلکه در عوض مواردی که در آنها همادی اکیداً تعریف‌پذیر وجود دارد استثنایی‌اند. این بدان معنی نیست که بگوییم آنها از نظر عددی اندک شمارند. بسیاری از آنها وجود دارند و واجد اهمیت ویژه‌اند. اما با این حال به نظر نمی‌رسد در کار عادی‌مان بخش بزرگتری از کاربست‌های استنباط محتمل را تشکیل دهند. کاربست‌های دیگر نیز به هیچ‌وجه پیش پا افتاده نیستند. کمبل[8] [۶]، در مقاله‌ای جالب که از جملهٔ سایر مطلب بخشی از زمینه‌ای را که تا کنون بررسی کرده‌ایم در بر دارد، به این نکته اشاره کرده است که تعریف بسامدی احتمال برای برقراری ارتباط بین مکانیک آماری و ترمودینامیک که یقیناً در نظریهٔ فیزیکی اهمیت اساسی دارد، کفایت نمی‌کند.

کینز[9] [۷]، با دقت و با زحمت فراوان روایتی بسیار اصیل از نظریهٔ احتمال عرضه کرده است که وابسته به مفهوم بسامد در یک هماد نیست. در دیدگاه وی، نظریهٔ احتمال منطقی گسترش یافته، منطق استنباط محتمل است. احتمال، رابطه‌ای بین یک فرضیه و یک نتیجه‌گیری است که متناظر با درجهٔ باور عقلانی و محدود به روابط کرانگین یقینی و عدم امکان است. منطق قیاسی کلاسیک که تنها به بررسی این روابط حدّی می‌پردازد حالتی خاص در محدودهٔ این ابداع کلی‌تر است. پس به طور کلی این نتیجه را در پی دارد که نظریهٔ احتمال نمی‌تواند کلاً بر پایهٔ مفهوم‌های منطق کلاسیک بنا نهاده شود. به‌ویژه، رابطه احتمال نمی‌تواند برحسب یقین تعریف شود، زیرا قطعیت خود حالتی خاص از احتمال است. بنابراین تعریف بسامدی احتمال نامعتبر است، زیرا به روابط یقینی ناظر به آگاهی از تعداد موارد وابسته است. احتمال به مثابه مفهومی نخستی، مانند مسافت یا زمان در مکانیک در نظر گرفته می‌شود که به هیچ نوع به مفهوم مقدماتی‌تر فروکاستنی نیست.

صرفاً توصیف موضع کینز بدون ذکر استدلالی که وی به این موضع رسیده، چنانکه من این کار را کردم، در خصوص کار او چندان عادلانه نیست. استدلال وی، دست کم برای من، بسیار

[8] E. C. Kemble, American Journal of Physics, ۱۰, 6 (1942).

[9] J. M. Keynes, A treaties on probability (Macmillan, London, 1929).

قانع کننده است و منبع اصلی بخش بزرگی از عقاید ابراز شده در اینجاست، هر چند که استدلال‌هایی که من به کار برده‌ام همانند استدلال‌های وی نیستند. با این وجود باید اذعان کرد که کار او از لحاظ حل مسئله‌ای که پیشتر ذکر شد، یعنی اثبات چند قاعدهٔ اساسی از استنباط محتمل که برای ابداع نظریهٔ مذکور لازم‌اند، ما را خیلی به جایی نمی‌رساند. این قاعده‌ها در نظریهٔ کینز صرفاً به مثابه اصل موضوع اختیار می‌شوند. اینک هر کس که دیدگاه کلی وی دربارهٔ ماهیت احتمال را می‌پذیرد، همان طور که من قویاً مایل به پذیرش آن هستم، باید چند فرض نخستی را بپذیرد، زیرا به محض اینکه تعریف بسامدی کنار گذاشته شد، نوعی نقطهٔ آغازین عقلانی لازم است که جانشین آن شود.اما اصل‌های موضوع کینز به نظر من، چنانچه بی‌تردید به نظر دیگران از جمله کمبل رسیده، تا اندازه‌ای بیش از حدّ دلبخواهی و بسیار پیچیده‌اند که به تمامی به مثابه اصل موضوع مناسب باشند. این اصل‌ها برای عقل سلیم به‌طور بی‌واسطه جالب نیستند، و درک اینکه بدون در نظر گرفتن گوی‌های رنگی در یک جعبه، تاس‌ها، سکه‌ها، یا برخی ابزارهای دیگر مرتبط با مفهوم هماد چگونه می‌توانسته‌اند فرمول‌بندی شوند، سخت است. چنین به نظر می‌رسد که گویی اقلیدس قضیهٔ فیثاغورس را جزو اصل‌های موضوع هندسهٔ هامنی جای داده است.

راسل[۱۰] [۱۰] انتقادی می‌کند که تا اندازه‌ای از لحاظ صورت متفاوت است اما ممکن است دلیلی مثل این انتقاد داشته باشد. با وجود این، او پس از اذعان به قوت استدلال کینز علیه نظریهٔ بسامدی، آن نظریه را به دلیل تعریف صریح آن از احتمال ترجیح می‌دهد به شرطی که بتواند به طور منطقی پایه‌گذاری شود. همین تعریف است که پرهیز از پیش‌فرض اصل‌های موضوعی را که نظریهٔ کینز را مشخص می‌سازند، ممکن می‌کند.

نویسندگان دیگری که مانند کینز، ابداع اصل موضوعی را عرضه می‌دارند، مجموعه‌هایی از اصل‌های مسلّم را برمی‌گزینند که تا حدّی متفاوت‌اند اما آنهایی که من دیده‌ام هنوز هم برخی از نشانه‌ها با خود دارند که اشتقاق اولیه‌شان از بررسی بازی‌های شانسی را، همراه با دلالت ضمنی بر وجود یک هماد، به نمایش می‌گذارند.

فکر می‌کنم این نکته حتی دربارهٔ اصل‌های مسلّمی که جفریز و رینج به دقت انتخاب کرده‌اند، خواه در صورت اولیه‌شان یا در صورت بازنگری شدهٔ آنها توسط جفریز[۱۱] [۵]، نیز صادق است.

۲ روابط انتظار معقول سازگار با منطق نمادین

در آنجه در زیر می‌آید، سعی خواهم کرد با به‌کارگیری جبر منطق نمادین نشان دهم که اشتقاق قاعده‌های احتمال از دو مفهوم نخستی که مستقل از مفهوم همادند میّسر است و به گمان من برای عقل سلیم بیشتر جذابیت دارند. این جبر از سوی شماری از نویسندگان، از جمله بول[۱۲] [۲]، که آن را به وجود آورد در احتمال به کار بسته شده است. با این حال، به نظر نمی‌رسد که امکانات آن در این خصوص کاملاً تحقق یافته باشند. به جا خواهد بود که در اینجا با جبر بولی، دست کم با قدری از آن که استدلال‌های بعدی لازم خواهند داشت، به کوتاهی آشنا شویم.

حرف‌های a، b، c، ... گزاره‌ها را نشان خواهند داد. سخن گفتن از احتمال‌های گزاره‌ها به جای احتمال‌های پیشامدها امتیازی دارد که تا حدودی به خاطر کلیت بیشتر اما عمدتاً به دلیل آن است که صحبت از پیشامدها بی‌تردید توالی بر حسب زمان را تداعی می‌کند، و این امر می‌تواند منبع سردرگمی شود. البته گزاره‌ای ممکن است رخداد پیشامدی را بیان کند اما همچنین می‌تواند چیزی دیگر، مثلاً چیزی دربارهٔ یک ثابت فیزیکی را، نشان دهد. گزارهٔ نه a با $\sim a$، گزارهٔ a و b با $a.b$، و گزارهٔ a یا b با $a \vee b$ نمایانده خواهد شد.

باید به خاطر داشت که گزارهٔ $\sim a$ گزاره‌ای خاص که به معنایی خلاف a باشد، نیست. بدین سان اگر a گزارهٔ، "شخص بیگانه مردی کوتاه قد، چاق، پیر بدون پالتو و کلاه بود" باشد، $\sim a$ گزارهٔ، "شخص بیگانه زنی بلند قد، لاغر، جوان با پالتو و کلاه بود" نیست. بیان $\sim a$ به معنای چیزی بیش از پاسخ "نه" به پرسش "آیا a کاملاً درست است؟" نیست. اگر a متشکل از چند پارهٔ a_1، a_2، ... باشد، بیان $\sim a$ اعلام این مطلب نیست که a_1، a_2، ... همه نادرست‌اند

[۱۰] B. Russell, Philosophy (Norton, New York, 1972).

[۱۱] H. Jefferys, Theory of probability (Clarendon Press, Oxford, 1939).

[۱۲] G. Boole, An investigation of the laws of thought (Macmillan, London, 1854).

بلکه این است که بگوییم دست کم یکی از آنها نادرست است.

چون حرف‌های a و b گزاره‌ها را می‌نمایانند و نه پیشامدها را، به ترتیبی که طی آن در نمادهای $a.b$ و $a \vee b$ ظاهر می‌شوند صرفاً همان ترتیبی است که دو گزاره بیان می‌شوند، نه ترتیبی که دو پیشامد در طی زمان رخ می‌دهند. همچنین صورت $a.a$ تنها بیانگر آن است که گزاره‌ای دو بار بیان شده است، نه آنکه پیشامدی دو بار رخ داده باشد.

همچنین باید فهمید که $a \vee b$ به معنای a یا b همان است که کودکی می‌پرسد، "آیا می‌توانم یک سکهٔ هزار ریالی یا یک سکه دو هزار ریالی داشته باشم؟" بدون آنکه منظورش حذف امکان داشتن هر دو سکهٔ هزار ریالی و دو هزار ریالی باشد، نه به آن معنا که خطیب می‌گوید "غرق شو یا شنا کن، زنده بمان یا بمیر." بدین ترتیب معنای $a \vee b$ همان است که اغلب به صورت a و / یا b به‌کار گرفته می‌شود.

سرانجام باید توجه داشت که اگر گزارهٔ، "باران می‌بارد" درست باشد، آن‌گاه گزارهٔ، "باران می‌بارد یا برف می‌بارد"، نیز درست است. اعلام گزارهٔ a دلالت بر هر گزارهٔ $a \vee b$ دارد که a جمله‌ای از آن است.

با درک معنای نمادها به ترتیبی که در بالا ذکر شدند، قاعده‌های ترکیب آنها را می‌توان به شرح زیر نمایش داد

$$\sim\sim a = a \tag{۱}$$

$$a.b = b.a \tag{۲}$$

$$a \vee b = b \vee a \tag{۲'}$$

$$a.a = a \tag{۳}$$

$$a \vee a = a \tag{۳'}$$

$$a.(b.c) = (a.b).c = a.b.c \tag{۴}$$

$$a \vee (b \vee c) = (a \vee b) \vee c = a \vee b \vee c \tag{۴'}$$

$$\sim (a.b) = \sim a \vee \sim b \tag{۵}$$

$$\sim (a \vee b) = \sim a . \sim b \tag{۵'}$$

$$a.(a \vee b) = a \tag{۶}$$

$$a \vee (a.b) = a \tag{۶'}$$

این یازده قاعده همه مستقل نیستند. این امکان هست که از شش تای آنها بقیهٔ پنج تای دیگر را اثبات کرد، و مجموعهٔ شش قاعده را می‌توان به راه‌های گوناگون انتخاب کرد. تنها لازم است که قاعدهٔ اول و یکی از هر یک زوج‌های دیگر را که شماره‌های مشابه دارند برگزید. بدین ترتیب، مثلاً رابطهٔ $(5')$ به ترتیب زیر به دست آورده می‌شود.

$$\begin{aligned} \sim (a \vee b) &= \sim (\sim\sim a \vee \sim\sim b) && [\text{بنابر رابطهٔ (۱)}] \\ &= \sim\sim (\sim a . \sim b) && [\text{بنابر رابطهٔ (۵)}] \\ &= \sim a . \sim b && [\text{بنابر رابطهٔ (۱)}] \end{aligned}$$

اکنون گیریم نماد $b|a$ ملاکی از باورمندی معقول گزارهٔ b برای وقتی باشد که می‌دانیم گزارهٔ a درست است[۱۳] [۹]. این دانشواژه در این مرحله نامعین است زیر اگر یک چنین ملاکی وجود داشته باشد تعداد زیادی از ملاک‌های دیگر نیز وجود خواهند داشت. اگر $b|a$ یک چنین ملاکی باشد، آن‌گاه تابع دلخواه $f(b|a)$ نیز یک ملاک خواهد بود. در نتیجه، لازم نیست اکنون این نماد با احتمال قراردادی یکی پنداشته شود. برای پرهیز از این دلالت، $b|a$ را درستنمایی گزارهٔ b تحت فرض a خواهیم نامید، تا از پیشنهادی که مارجینو[۱۴] [۸] کرده بهره بگیرم اما در همان حال آزادی عمل داشته باشم که به این دانشواژه معنایی فراگیرتر از معنای پیشنهادی ببخشم.

نباید گمان کرد که رابطهٔ درستنمایی بین هر دو گزاره وجود دارد. اگر a گزارهٔ "سزار به بریتانیا حمله کرد" و b "فردا گرمتر از امروز خواهد بود" باشد، درستنمایی $b|a$ وجود ندارد، زیرا هیچ رابطهٔ معقول بین این دو گزاره برقرار نیست.

اکنون زمان آن فرا رسیده است که پیش فرض اول از دو پیش فرضی را که پیشتر به عنوان فراهم آورندهٔ پایه‌ای برای اصول استنباط محتمل ذکر شدند، بیان کنیم. هر ملاکی که انتخاب شده باشد، فرض خواهیم کرد درستنمایی $c.b|a$ به طریقی به وسیلهٔ دو

[۱۳] کینز استفاده از این نماد را تا کار H. McColl, Proc. Lond. Math. Soc. 11, 113 (1880). پیگیری کرده است.مه‌کال نماد x_a را برای احتمال گزارهٔ x تحت فرضیه a به کار می‌برد.

[۱۴] H. Margenau, American Journal of Physics 10, 224 (1942). آر. ای. فیشر این دانشواژه را به معنای کاملاً متفاوت به کار برده است.

درستنمایی $b|a$ و $c|b.a$، یا

$$c.b|a = F(c|b.a, b|a) \qquad (۷)$$

تعیین می‌شود که در آن F تابعی از دو متغیر است. این پیش فرض که به صورت نمادین نوشته شده، ممکن است خیلی اصل موضوعی به نظر نرسد. درواقع، چنانکه مثالی نشان خواهد داد، این قاعده‌ای بسیار آشنا از عقل سلیم است. گیریم b گزاره‌ای را بنمایاند که دونده‌ای می‌تواند از یک جای معین به جای دیگر بدود، و گیریم c این گزاره باشد که او می‌تواند بدون توقف دویده به جای اول برگردد. شرایط فیزیکی و عارضه نگاری مسیر دو در فرضیهٔ a توصیف می‌شوند. در این صورت $b|a$ درستنمایی برآورد شده بر پایهٔ اطلاعات موجود در a است که او می‌تواند تا آن مکان دور بدود، و $c|b.a$ درستنمایی برآورد شده بر پایهٔ اطلاعات آغازین و پیش فرض دیگر است که وی هم اکنون یک راه را دویده و می‌تواند مسیر برگشت را بدود. اینها تنها درستنمایی‌هایی هستند که برای برآورد درستنمایی $c.b|a$، یعنی اینکه او می‌تواند مسیر کامل را بدون توقف بدود، باید در نظر گرفته شوند. با مسلّم دانستن اینکه درستنمایی یاد شدهٔ اخیر تابعی از دودرستنمایی دیگر است، کم قیدترین پیش فرض ممکن را می‌پذیریم.

به دلیل نامعین بودن ملاکی که باید برای درستنمایی به کار رود، ضابطهٔ تابع F تا حدی قراردادی است. اما به طور کامل چنین نیست، زیرا باید با جبر گزاره‌ها سازگار باشد. بنابراین از رابطهٔ (۴) برای به دست آوردن یک معادلهٔ تابعی شامل F به شرح زیر استفاده می‌کنیم:

$$\begin{aligned} d.c.b|a &= (d.c).b|a \quad \text{[بنابر رابطهٔ (۴)]} \\ &= F(d.c|b.a, b|a) \quad \text{[بنابر رابطهٔ (۷)]} \end{aligned}$$

اما

$$\begin{aligned} d.c|b.a &= F(d|c.(b.a), c|b.a) \quad \text{[بنابر رابطهٔ (۷)]} \\ &= F(d|c.b.a, c|b.a) \quad \text{[بنابر رابطهٔ (۴)]} \end{aligned}$$

بنابراین

$$d.c.b|a = F[F(d|c.b.a, c|b.a), b|a]$$

همچنین

$$\begin{aligned} d.c.b|a &= d.(c.b)|a \quad \text{[بنابر رابطهٔ (۴)]} \\ &= F[d|(c.b).a, c.b|a] \quad \text{[بنابر رابطهٔ (۷)]} \\ &= F[d|c.b.a, F(c|b.a, b|a)] \quad \text{[بنابر رابطهٔ (۴) و (۷)]} \end{aligned}$$

پس از برابر نهادن این دو عبارت مربوط به $d.c.b|a$ و برای سادگی کار با قرار دادن $d|c.b.a = x$، $c|b.a = y$، و $b|a = z$، داریم

$$F[F(x,y), z] = F[x, F(y,z)] \qquad (۸)$$

تابع F باید چنان باشد که به‌ازای مقدارهای دلخواه x، y، و z در معادلهٔ (۸) صدق کند. با انجام جایگذاری به آسانی نشان داده می‌شود که این معادله در صورتی صادق است که داشته باشیم

$$Cf[F(p,q)] = f(p)f(q)$$

که در آن f تابعی دلخواه از متغیری تک، و C ثابتی دلخواه است. در پیوست نشان داده می‌شود که این جواب کلی نیز هست، به شرط آنکه F دارای مشتق‌های دوم باشد. در آن صورت داریم

$$Cf(c.b|a) = f(c|b.a)f(b|a)$$

انتخاب تابع f صرفاً امری قراردادی است. زیرا پیشتر اشاره شد که اگر $b|a$ ملاکی از باورمندی b تحت فرض a باشد، آن‌گاه $f(b|a)$ نیز چنان است. دراین صورت می‌توانیم به جای استفاده از $b|a$ بحث را با $f(b|a)$ به عنوان نماد درستنمایی ادامه دهیم و هرگز ناگزیر نباشیم تابع f را مشخص کنیم. اما این کار دو نماد خواهد داد در جایی که یک نماد کافی است. بنابراین، برای راحتی کار می‌نویسیم

$$C\ c.b|a = c|b.a\ b|a \qquad (۹)$$

البته این کار مثل انتخاب تابع f است که $f(b|a) = b|a$ را پدید آورد. چون انتخاب قراردادی بود، نتیجه می‌شود که انتخاب دیگری نیز می‌شد به عمل آورد. برای مثال، ممکن بود $f(b|a) = \exp(b|a)$ را اختیار کنیم که از آن نتیجه می‌شد درستنمایی $c.b$ به جز یک ثابت دلخواه برابر با مجموع درستنمایی‌هایی است که آن را تعیین می‌کنند. این کار یک درستنمایی مرتبط با نوعی درستنمایی را به ما می‌داد که به صورت آنتروپی مرتبط با احتمال ترمودینامیک در مکانیک آماری داریم. این تابع گزینه‌ای مجاز خواهد بود اما گزینه‌ای راحت‌تر از آنی که انتخاب کردیم نیست. در رابطهٔ (۹) قرار می‌دهیم $c = b$ و دقت می‌کنیم که بنابر رابطهٔ (۳)، $b.b = b$، پس از تقسیم بر $b|a$ به دست می‌آوریم

$$C = b|b.a$$

بدین ترتیب، می‌بینیم که وقتی گزارهٔ فرضیه گزارهٔ نتیجه‌گیری را دربرداشته باشد، گزاره‌ها هر چه باشند، درستنمایی مقدار ثابت C را دارد. این حکم چیزی است که باید انتظار داشته باشیم، زیرا تحت فرضیهٔ $a.b$، گزارهٔ b یقینی است و دیگر درجه‌های یقین را به رسمیت نمی‌شناسیم.

مقداری که باید به C، یعنی درستنمایی یقین نسبت داده شود، کاملاً قراردادی است. اگر خواسته باشیم درستنمایی‌هایی که با آنها سروکار داریم تا حد امکان نزدیک به احتمال‌های عادی باشند، آن‌گاه مقدار یک را به C نسبت خواهیم داد. گزینه‌های دیگر، به ویژه در گفتگوها، انتخاب می‌شوند. عبارت "یک شانس در یک صد شانس" را می‌توان به معنای درستنمایی واحد در مقیاسی گرفت که در آن یقین با ۱۰۰ نمایش داده می‌شود. حکم‌هایی که به صورت ادعاهایی دربارهٔ اعداد موجود در یک هماد بیان می‌شوند ممکن است راه‌های صرفاً آسان برای بیان درستنمایی‌ها در مقیاسی باشند که به خاطر تناسب آن با سوال مورد نظر انخاب شده است. در یک بحث کلی، راحت‌ترین مقدار برای C عدد یک است، و از این رو رابطهٔ (۹) را به صورت زیر می‌نویسیم

$$c.b|a = c|b.a.b|a \qquad (۱۰)$$

این رابطه مشابه با صورت قاعدهٔ عادی مربوط به احتمال دو پیشامد است. اما باعث نمی‌شود که درستنمایی ما به طور یکتا با احتمال عادی برابر باشد. زیرا رابطهٔ (۱۰) به هر مرتبه‌ای از توان m که برسد عبارت است از

$$(c.b|a)^m = (c|b.a)^m (b|a)^m$$

بدین ترتیب، هر توانی از درستنمایی ما از حیث صورت در رابطه‌ای مشابه با رابطهٔ (۱۰) صدق می‌کند و به همان اندازه با احتمال عادی متناظر است.

چیزی که در مرحلهٔ بعد باید جستجو شود یک پیش فرض دوم برای استنباط محتمل است که قرار است رابطه‌ای را بین درستنمایی‌های گزاره‌های b و $\sim b$ تحت فرضیه واحد a فراهم سازد. چون وقتی b مشخص باشد، $\sim b$ تعریف شده است، پیش فرضی معقول و با امکان قید کمتر به نظر می‌رسد پیش فرضی باشد که $\sim b|a$ به وسیلهٔ $b|a$ تعیین شود، یا

$$\sim b|a = S(b|a) \qquad (۱۱)$$

باشد که در آن S تابعی از متغیری تک است.

بنابر رابطهٔ (۱)، $b|a = b|a \sim\sim$، و بنابراین $S[S(b|a)] = b|a$. پس S باید چنان تابعی باشد که داشته باشیم

$$S[S(x)] = x \qquad (۱۲)$$

که در آن x می‌تواند هر مقدار ممکن از درستنمایی را بین درستنمایی‌های یقین و ناممکن داشته باشد. این خاصیت قید کافی بر S تحمیل نمی‌کند که به تنهایی فایدهٔ زیادی داشته باشد. معادلهٔ تابعی دیگر را می‌توان با در نظر گرفتن $S(c \vee b|a)$ به دست آورد؛ بدین سان،

$$S(c \vee b|a) = \sim (c \vee b|a) = [(5') \text{ بنابر رابطهٔ }] \sim c. \sim b|a.$$

مطلوب آن است که گزاره‌های $\sim c$ و $\sim b$ را حذف کنیم، به طوری که معادله‌ای بر حسب گزاره‌های c، b، و a و تابع S به دست آوریم. نخست $\sim c$ را حذف می‌کنیم.

$$\begin{aligned} \sim c. \sim b|a &= [(۱۰) \text{ بنابر رابطهٔ }] \sim c| \sim b.a \sim b|a \\ &= [(۱۱) \text{ بنابر رابطهٔ }] S(c| \sim b.a)S(b|a). \end{aligned}$$

پس داریم

$$S(c \vee b|a) = S(c| \sim b.a) - S(b|a)$$

یا

$$S(c| \sim b.a) = S(c \vee b|a)S(b|a)$$

در دو طرف رابطهٔ بالا تابع S را به کار برده از رابطهٔ (۱۲) استفاده می‌کنیم، و به دست می‌آوریم

$$c| \sim b.a = S[S(c \vee b|a)/S(b|a)] \qquad (۱۳)$$

در مرحلهٔ بعد b را حذف می‌کنیم:

$$\begin{aligned} c| \sim b.a &= [(۱۰) \text{ بنابر رابطهٔ }] c. \sim b|a/ \sim b|a \\ &= [(۲) \text{ بنابر رابطهٔ }] \sim b.c|a/ \sim b|a \\ &= [(۱۰) \text{ بنابر رابطهٔ }] \sim b|c.a\ c|a/ \sim b|a \\ &= [(۱۱) \text{ بنابر رابطهٔ }] S(b|c.a)\ c|a/S(b|a) \end{aligned}$$

بنابراین به جای رابطهٔ (۱۳) می‌توانیم بنویسیم

$$S(b|c.a)c|a/S(b|a) = S[S(c \vee b|a)/S(b|a)].$$

راحت‌تر است که a را به عنوان فرضیه مشترک در همهٔ درستنمایی‌ها داشته باشیم. توجه داریم که

$$b|c.a = [(۱۰) \text{ بنابر رابطهٔ }] b.c|a/c|a$$

$$= \quad [\text{بنابر رابطهٔ (۲)}]\, c.b|a/c|a$$

پس از جایگذاری این عبارت در رابطهٔ قبلی و ضرب کردن دو طرف در $S(b|a)$، به دست می‌آوریم

$$S(c.b|a/c|a)c|a = S[S(c \vee b|a)/S(b|a)]S(b|a) \qquad (۱٤)$$

این رابطه باید به ازای معانی دلخواه گزاره‌های a، b، و c برقرار باشد. گیریم $b = c.d$. آن‌گاه

$$c \vee b = c \vee (c.d) = [\text{بنابر رابطهٔ } (6')]\, c$$

و

$$c.b = c.(c.d) = [\text{بنابر رابطهٔ (۱٤)}]\,(c.c).d = [\text{بنابر رابطهٔ (۳)}]\, c.d$$

با جایگذاری این عبارت‌ها در رابطهٔ (۱٤) به دست می‌آوریم

$$S(c.d|a/c|a)c|a = S[S(c|a)/S(c.d|a)]S(c.d|a)$$

اگر قرار دهیم $c|a = x$ و $S(c.d|a) = y$ و از این واقعیت که

$$c.b|a = [\text{بنابر رابطهٔ (۱۲)}]\, S[S(c.d|a)] = S(y)$$

بهره گیریم، این رابطه می‌تواند به صورتی بسیار متقارن نوشته شود. برحسب جمله‌های x و y، داریم

$$xS[S(y/x)] = yS[S(x/y)] \qquad (۱۵)$$

این معادله باید به ازای همهٔ مقدارهای x و y که از راه تغییر دادن دلبخواهی گزاره‌های c، d، و a به دست آمدنی‌اند برای تابع S برقرار باشد. اگر تابع S دو بار مشتق‌پذیر باشد، جواب معادلهٔ (۱۵) همراه با رابطهٔ (۱۲)، چنانکه در پیوست نشان داده شده، عبارت است از

$$S(p) = (1 - p^m)^{\frac{1}{m}}$$

که در آن m ثابتی دلخواه است. بنابراین، طبق رابطهٔ (۱۱)،

$$(b|a)^m + (\sim b|a)^m = 1$$

اینک، مقدار m هر چه باشد، اگر $b|a$ باورمندی b تحت فرضیه a را بسنجد، $(b|a)^m$ نیز چنان خواهد کرد. پیشاپیش اشاره شده است که $(b|a)^m$ می‌تواند جانشین $b|a$ در رابطهٔ (۱۰) شود. بنابراین می‌توانیم $(b|a)^m$ را به مثابه نماد درستنمایی بگیریم بدون آنکه با هیچ ضرورتی به انتساب مقداری به m مواجه باشیم. این گفته معادل با آن است که بگوییم انتخاب مقداری برای m کاملاً قراردادی است. برای سادگی نمادگذاری قرار می‌دهیم $m = 1$ و می‌نویسیم

$$b|a + \sim b|a = 1 \qquad (١٦)$$

این رابطه همان صورتی را دارد که قاعدهٔ عادی احتمال $\sim b$ را به احتمال b ربط می‌دهد، یا چنانکه معمولاً گفته می‌شود، قاعدهٔ مربوط به احتمال زخ ندادن پیشامدی به شرط معلوم بودن احتمال آنکه آن پیشامد رخ خواهد داد.

در رابطهٔ (١٦) قرار می‌دهیم $b = a$، آن‌گاه

$$a|a + \sim a|a = 1.$$

این دو درستنمایی اکنون درستنمایی‌های یقین و ناممکن‌اند. چون درستنمایی ۱ به گزارهٔ یقین نسبت داده شده است، نتیجه می‌شود که ناممکن دارای درستنمایی صفر است.

دو قضیهٔ سودمند دیگر به آسانی از رابطهٔ (۱۰) به دست می‌آیند،

$$c.b|a + \sim c.b|a = (c|b.a + \sim c|b.a)b|a$$

بنابر رابطهٔ (١٦)،

$$c|b.a + \sim c|b.a = 1$$

بنابراین

$$c.b|a + \sim c.b|a = b|a \qquad (۱۷)$$

این یکی از قضیه‌هاست. دیگری به شرح زیر به دست می‌آید.

$$\begin{aligned} c \vee b|a &= [\text{بنابر رابطهٔ (١٦)}]\, 1 - \sim (c \vee b)|a \\ &= [\text{بنابر رابطهٔ } (5')]\, 1 - \sim c. \sim b|a \\ &= [\text{بنابر رابطهٔ (۱۷)}]\, 1 - \sim b|a + c. \sim b|a \end{aligned}$$

اکنون، طبق رابطهٔ (١٦)، داریم $1 - \sim b|a = b|a$. همچنین

$$\begin{aligned} c. \sim b|a &= [\text{بنابر رابطهٔ (۲)}] \sim b.c|a \\ &= [\text{بنابر رابطهٔ (۱۷)}]\, c|a - b.c|a \\ &= [\text{بنابر رابطهٔ (۲)}]\, c|a - c.b|a \end{aligned}$$

بنابراین،

$$c \vee b|a = c|a + b|a - c.b|a \qquad (۱۸)$$

که همان صورت قاعدهٔ عادی را برای احتمال آنکه دست کم یکی از دو پیشامد رخ بدهد، داراست. چنانکه اغلب به کار می‌رود، وقتی این قاعده برای پیشامدهای دو به دو ناسازگار بیان شود، جملهٔ آخر

در سمت راست رابطهٔ بالا ظاهر نخواهد شد. این امر تقارن نسبتاً جالب رابطه را بین گزاره‌های $c \vee b$ و $c.b$ پوشیده نگاه می‌دارد.

نامعین بودن مفهوم درستنمایی که تنها به مثابه ملاکی معقول از باورمندی تعریف شده بود، به کمک قراردادهایی که برگزیده شده‌اند حذف شده است. نماد $b|a$ اکنون نشانهٔ ملاکی ویژه از باورمندی است. چون نشان داده شده است که این ملاک مشمول قاعده‌های عادی احتمال است، مقتضی است آن را احتمال گزارهٔ b تحت فرضیهٔ a نامیده، دانشواژهٔ درستنمایی را که با دقت کمتر تعریف شده بود کنار بگذاریم.

قاعده‌هایی که به دست آورده شده‌اند تنها روابط بین احتمال‌ها هستند و به خودی خود مقدارهای عددی را به همهٔ احتمال‌هایی که در مسئله‌های خاص ظاهر می‌شوند نسبت نمی‌دهند. تنها مقدارهای عددی که تا کنون به دست آورده شده‌اند عددهایی‌اند که متناظر با یقین و ناممکن‌اند، و این مقدارها به جای آنکه بنابر قاعده‌های منطق نمادین لازم باشند، طبق قرارداد نسبت داده شدند. به سختی می‌توان گمان برد که هر انتظار معقول باید یک مقدار عددی دقیق داشته باشد. اما در شماری از موارد، قاعدهٔ آشنای دلیل ناکافی را می‌توان به کار گرفت. اگر n گزاره وجود داشته باشند که با توجه به فرضیه‌ای معین یکی و نه بیش از یکی از آنها تواند درست بود، و اگر این فرضیه دلیلی برای در نظر گرفتن هر یکی از آنها محتمل‌تر از دیگری در اختیار نگذارد، آن‌گاه طبق قاعده‌های به دست آمده، هر یک از آنها احتمال $\frac{1}{n}$ خواهد داشت.

۳ احتمال و بسامد

کل بحث تا کنون مشتمل بر دو بخش بوده است. مراد از بخش نخست این بود که نشان دهیم قاعده‌های استنباط محتمل از اعتبار عقل سلیم برخوردارند، با اعتباری گسترده‌تر از آنچه بتوان با اشتقاق آنها از تعریف بسامدی احتمال ثابت کرد. در بخش دوم این قاعده‌ها بدون ارجاع به این تعریف، از اصل‌های مسلّم نسبتاً ابتدایی به دست آورده شدند. اینک چیزی که باقی می‌ماند، درک ارتباط بین احتمال، به معنایی که در اینجا فهمیده می‌شود، و بسامد یک پیشامد است.

گیریم فرض کنیم که دو پوشینه حاوی جرم‌هایی برابر از رادون، اما محتوای پوشینه‌ها گازهای متفاوت باشند، یکی از تلاشی بسیار جدید رادیوم تولید شده، دیگری از ظرفی استخراج شده که در آن رادون به مدت طولانی روی رادیوم درون مایع انباشته شده است. فرض کنید دو شمارشگر یون همانند وجود دارند که هر یک از آنها یک پوشینه را دریافت می‌کند و هر یک با توجه به پوشینه‌اش نسبت به دیگری در همان موضع نسبی قرار داده می‌شود. یکی از دو پوشینه نخستین پوشینه‌ای خواهد بود که موجب ۱۰۰۰ تخلیه در شمارشگر یون خود خواهد شد. بیش از یک فرضیه به هر پوشینه احتمال‌های مساوی مبنی بر اول بودن را نسبت خواهد داد. یک فیزیکدان به دلیل مشاهدات پرشماری که از نرخ‌های تبدیل پرتوزایی به عمل آمده‌اند، با قدری شواهد اضافی از مکانیک کوانتومی مبنی بر اینکه پایداری چنین انبوهه‌ای از ذرات ابتدایی به عنوان هستهٔ اتمی مستقل از سن آن است، احتمال‌های برابر را برآورد خواهد کرد. شخصی دیگر، کاملاً ناآشنا به همهٔ این مطالب، احتمال‌های برابر را صرفاً به دلیل اینکه نمی‌داند کدام پوشینه رادون مسن‌تر را دربردارد و بنابراین هیچ دلیل ممکنی ندارد که گمان کند یک نمونه نسبت به نمونهٔ دیگر نخستین نمونه خواهد بود که باعث ۱۰۰۰ تحلیل خواهد شد، برآورد خواهد کرد.

این برآوردها حالت‌های کرانگین‌اند، برآورد اول به شدت معنی‌دار و دومی کاملاً پیش پا افتاده است، اما تحت فرضیه داده شده هر یک از انها درست است. کمبل برآورد فیزیکدان را از نوع احتمال عینی و از آن شخص دیگر را از نوع احتمال ذهنی، یا اولیّه می‌نامد. دانشواژه آخر برایم مرجح به نظر می‌رسد، چنانکه برای وی نیز چنین است. درست است که برآورد غیر فیزیکدان ذهنی است به این معنا که با اطلاعات محدود وی مرتبط است، اما عینی است به این معنا که شخصی دیگر نیز با هیمن اطلاعات به طور معقول همین برآورد را خواهد کرد. به نظر قابل تردید می‌رسد که تفاوتی حقیقی در نوع داوری‌یی که غیر فیزیکدان و فیزیکدان به عمل می‌آورند وجود دارد یا نه. غیر فیزیکدان برآوردش را بر پایهٔ این واقعیت که پوشینه‌ها غیر قابل تمیزند بنا می‌نهد. فیزیکدان برآوردش را بر پایهٔ شواهد انباشته شده که اتم‌های رادون خودشان غیرقابل تمیزند بنا می‌گذارد. به نظر می‌رسد که این تفاوت از حیث ماهیت شواهد به لحاظ مقدار و مناسبت آن با مسئله، یا با استفاده از دانشواژهٔ دلالت کننده کینز، از لحاظ وزن آن چندان قابل توجه نیست.

اکنون گیریم آزمایش با پوشینه‌های رادون چندبار انجام یابد که در آنها نمونه‌های کهنه و نو پیش از هر آزمایه شناسایی می‌شوند. حتی سلسله‌ای دراز از موارد که در آنها نمونهٔ مسن‌تر در جای اول است احتمال‌ها را آن‌گونه که فیزیکدان برآورد کرده، تغییر نخواهد داد. شواهدی که وی براساس آنها برآورد اولش را انجام داد چنان وزنی دارند که هیچ شمار اضافی، نه فوق‌العاده بزرگ، از موارد می‌توانست برآوردی جدید را لازم بنماید. احتمال‌ها برای هدف‌های عملی پایدار شده‌اند. دقیق‌تر بگوییم، چون احتمال مرتبط با تجربه‌ای است که هرگز کامل نیست، همواره مستعد تغییر با تجربهٔ جدید است. احتمال پایدار حدی است که به طور اکید دست یافتنی نیست، اما در موارد معین می‌تواند تا حد ضرورت برای کاربست عملی دقیق‌تر تقریب زده شود. می‌توان انتظار داشت که یک احتمال پایدار بهتر از یک احتمال ناپایدار، مبنایی برای پیشگویی در اختیار خواهد گذاشت.

گیریم a فرضیه‌ای باشد که با آن شماری از موارد را می‌توان امتحان کرد. گیریم b_r به معنای آن باشد که گزارهٔ معین b در مورد rام از a معتبر است. آن‌گاه $b_s|a$، $b_s|a.b_r$ و $b_s|a. \sim b_r$ به طور کلی متفاوت خواهند بود، مگر آنکه a خود احتمالی پایدار را به b نسبت دهد؛ آگاهی از اینکه b در یک مورد معتبر است یا معتبر نیست بر انتظار معقول از اعتبار آن در موردی دیگر تأثیر خواهد داشت. اما کنون گیریم گذارهٔ P در فرضیه‌ای گنجانده شود، که ادعا دارد احتمال پایدار و برابر p، عددی بین ۰ و ۱ است. این امر بدان معناست که خواه b_r یا $\sim b_r$ یا هیچ‌یک از آن دو در فرضیه گنجانده شده باشد یا نه، احتمال b_r یک مقدار معین است. بدین ترتیب

$$b_s|a.P.b_r = b_s|a.P. \sim b_r = b_s|a.P = p$$

آن‌گاه بنابر رابطه‌های (۱۰) و (۱۶) به دست می‌آوریم

$$\begin{aligned} b_s.b_r|a.P &= p^2 \\ b_s. \sim b_r|a.P &= p(1-p) \\ \sim b_s. \sim b_r|a.P &= (1-p)^2 \end{aligned}$$

گیریم n_N به معنای آن باشد که شمار موارد b در N مورد از a دقیقاً برابر n است. آن‌گاه بنابر رابطه‌های (۱۰)، (۱۶)، و (۱۸)، به دست آوردن نتیجهٔ مشهور برنولی میّسر می‌شود، که

$$n_N|a.P = p^n(1-p)^{N-n}\frac{N!}{n!(N-n)!}$$

این مقدار وقتی $p = \frac{n}{N}$ باشد، ماکسیمم است و به تدریج که N افزایش می‌یابد، ماکسیمم دقیق‌تر می‌شود. بدین سان، هنگامی که یک احتمال پایدار وجود داشته باشد، با اطمینان می‌توان انتظار داشت که بسامد به آن احتمال به عنوان یک حد بگراید.

گاهی پرسش‌هایی مطرح خواهند شد که در آنها وجود یک احتمال پایدار معلوم بوده اما مقدار آن نامعین است. به عنوان یک مثال نسبتاً تصنّعی اما ساده گیریم فرض شود که دو تاس وجود دارند، هر دو از لحاظ دینامیکی متقارن‌اند، اما یکی از آنها به طور معیوب شماره گذاری شده به جای یک وجه با چهار نقطه دو روی چهار نقطه‌ای دارد. در این صورت برای هر یک از این دو تاس یک احتمال پایدار از آمدن چهار وجود دارد، که اگر تاس درست باشد برابر $\frac{1}{6}$ و اگر تاس معیوب باشد برابر $\frac{1}{3}$ است. فرض کنید یک تاس از این زوج به تصادف انتخاب شده و بدون معاینهٔ آن، N بار پرتاب می‌شود. اگر در n تا از این پرتاب‌ها روی چهار ظاهر شود، احتمال آمدن چهار در پرتاب بعدی چیست؟

این مسئله را به شرح زیر می‌توان تعمیم داد. گیریم فرض شود که در همادی از موارد گزاره‌ای مانند a، معلوم است که دیگر گزاره‌ای مانند b احتمالی پایدار دارد اما مقدار این احتمال پایدار مجهول است. مانند قبل، P گزاره‌ای را خواهد نمایاند که احتمال پایدار بوده برابر عدد p است، اما در این حالت P بخشی از فرضیه نیست. بلکه فرضیه مذکور گزاره‌ای ضعیف‌تر را دربر دارد که تنها احتمالی را به گزارهٔ P متناظر با هر مقدار از p، نسبت می‌دهد. می‌توانیم این‌گونه فرض کنیم که تک نماد a کل فرضیه آغازین، از جمله این گزاره را نمایش دهد. بدین ترتیب، در مثال تاس، a دو تاس را توصیف خواهد کرد و همچنین بیان خواهد کرد که یکی به تصادف انتخاب شده بدون معاینه پرتاب می‌شود. در این صورت تنها دو احتمال پایدار $\frac{1}{6}$ و $\frac{1}{3}$ وجود دارند، و در آغاز به یک اندازه محتمل‌اند. بنابراین در این مثال $p|a$ مقدار $\frac{1}{2}$ را دارد اگر p $\frac{1}{6}$ یا $\frac{1}{3}$ باشد و مقدار صفر را دراد اگر p هر عددی دیگر باشد.

به مسئله کلی برگردیم، فرض می‌کنیم که N مورد از a مشاهده می‌شوند، و b در n تا از آنها معتبر و در بقیه نامعتبر یافت می‌شود. اکنون احتمال b در مورد $N+1$ام از a چیست؟ مثل قبل، گیریم

n_N این گزاره را بنمایاند که n شمار موارد b در N مورد از a است. مسئله پیدا کردن $b_{N+1}|a.n_N$ به شرط معلوم بودن n و N و نیز $P|a$ به ازای هر مقدار از p است.

قضیه‌های موجود کافی‌اند که نتیجهٔ زیر را به دست دهند

$$b_{N+1}|a.n_N = \frac{\sum p^{n+1}(1-p)^{N-n}p|a}{\sum p^n(1-p)^{N-n}p|a}$$

که در آن مجموع‌یابی‌ها نسبت به همهٔ مقدارهای p انجام می‌گیرند.

اگر تحت فرضیه a، احتمال پایدار دامنه‌ای پیوسته از مقداهای ممکن ۰ تا ۱ را داشته باشد، و اگر $f(p)dp$ احتمال مقداری بین p و $p+dp$ را نشان دهد، به جای مجموع‌ها انتگرال‌ها می‌نشینند، و داریم

$$b_{N+1}|a.n_N = \frac{\int_0^1 p^{n+1}(1-p)^{N-n}f(p)dp}{\int_0^1 p^n(1-p)^{N-n}f(p)dp}$$

لاپلاس به عنوان پیش‌فرض پذیرفت که احتمال مجهول به یک اندازه محتمل است که هر مقداری بین ۰ و ۱ را داشته باشد. تحت این پیش فرض، $f(p)$ در رابطهٔ اخیر مقداری ثابت است. در این حالت انتگرال‌ها معلوم‌اند؛ و نتیجه که گاهی قاعدهٔ توالی نامیده می‌شود صرفاً برابر است با

$$b_{N+1}|a.n_N = \frac{(n+1)}{(N+2)}$$

یا به ازای مقدارهای بزرگ از n و N، تقریباً برابر است با

$$b_{N+1}|a.n_N = \frac{n}{N}$$

چند مؤلف خاطر نشان کرده‌اند که احتمال پایدار مجهول لزوماً احتمالی نیست که همهٔ مقدارهایش از ۰ تا ۱ یه یک اندازه محتمل باشند، و نشان داده شده است که قاعدهٔ توالی به نتیجه‌هایی آشکارا نامعقول می‌انجامد. با وجود این، باید انتظار داشته باشیم که به ازای عددهای بزرگ این قاعده به طور کلی درست خواهد بود. زیرا طبق قاعدهٔ برنولی که پیشتر داده شد، می‌دانیم زمانی که یک احتمال پایدار وجود دارد، وقتی N بزرگ باشد، نسبت $\frac{n}{N}$ خیلی محتمل است که تقریباً برابر آن احتمال باشد. همچنین فهمیده‌ایم، احتمال پایدار حدی است که وقتی وزن شواهد افزایش یابد احتمال به آن می‌گراید، و معمولاً مطمئن‌ترین راه افزایش وزن شواهد افزایش شمار موارد مشاهده شده است. اگر نسبت $\frac{n}{N}$ و احتمال به حدی مشترک بگرایند، یقیناً آن دو باید به یک‌دیگر می‌گرایند.

اگر موارد نامعقولی که قاعدهٔ لاپلاس به آنها انجامیده بررسی شوند، معلوم خواهد شد که آنها به یکی از سه رده تعلق دارند: مواردی که در آنها N عددی خیلی بزرگ نیست، مواردی که در آنها $\frac{n}{N} = 1$ و مواردی که در آنها $\frac{n}{N} = 0$. (دو ردهٔ اخیر در حقیقت یک رده‌اند زیرا گفتن اینکه b در N تا از N مورد a معتبر است مثل این است که گفته شود $\sim b$ در هیچ یک از N مورد معتبر نیست). اگر این شرایط کنار گذاشته شوند، قاعدهٔ لاپلاس می‌تواند از پیش فرضی ضعیف‌تر از پیش فرض مبنی بر اینکه همه مقدارهای احتمال پایدار به یک اندازه محتمل‌اند، به دست آورده شود.

اگر از رابطهٔ کلی مربوط به $b_{N+1}|a.n_N$ با قرار دادن $\frac{n}{N} = \nu$، عدد n را حذف کنیم، به دست می‌آوریم

$$b_{N+1}|a.n_N = \frac{\int_0^1 p[p^\nu(1-p)^{1-\nu}]^N f(p)dp}{\int_0^1 [p^\nu(1-p)^{1-\nu}]^N f(p)dp}$$

اگر $1 > \nu > 0$ باشد، آن‌گاه $p^\nu(1-p)^{1-\nu}$ مقدار ماکسیمم را به ازای $p=\nu$ دارد. توان Nام این عبارت، وقتی N به قدر کافی بزرگ باشد، دارای ماکسیممی چنان برجسته خواهد بود که اگر p بیش از اندکی تفاوت با ν داشته باشد، مقدارهای آن به طور نسبی قابل چشم پوشی خواهند بود. از این رو، به جز حالتی که $f(p)$ در $p=\nu$ فوق‌العاده کوچک باشد، تنها مقدارهای واجد اهمیت آن در انتگرال‌ها، آن مقدارهایی خواهند بود که به ازای آنها p و ν بسیار نزدیک به هم‌اند. بنابراین، به جز از حالتی که $f(p)$ پیرامون این نقطه دارای تغییرات تند باشد، می‌توان به جای آن در انتگرال‌ها $f(\nu)$ی ثابت را جانشین کرد.

بدین ترتیب دوباره به قاعدهٔ لاپلاس می‌رسیم. کلیت آن بسیار کمتر از آن است که لاپلاس گمان می‌برد. اما برای نشان دادن اینکه وقتی شمار موارد افزایش می‌یابد، چگونه احتمال به پایداری می‌گراید به کار می‌رود، و این همهٔ آن چیزی است که باید از آن انتظار داشته باشیم[۱۵] [۲].

[۱۵] مسئله احتمال وارون هنگامی که $\frac{n}{N} = 1$ (یا ۰)، که در کاربرد احتمال در استدلال استقرایی اهمیت دارد، به تفصیل توسط جفریز در مرجع ۸ بحث شده است.

تقدیر و تشکر

استاد ک. او. فریدریکز از دانشگاه نیویورک پیش‌نویس مقدماتی این مقاله را خوانده است. مایلم به خاطر این محبت و به خاطر یاری‌اش در تصحیح برخی نادرستی‌های ریاضی از وی سپاسگزاری کنم. البته، وی برای هر گونه خطایی که باقی مانده باشد یا برای عقیده‌ای که دربارهٔ ماهیت احتمال ابراز شده است، مسئول نیست.

پیوست: حل معادله‌های تابعی

نخستین معادله‌ای که باید حل شود عبارت است از

$$F[F(x,y),z]=F[x,F(x,z)] \qquad (۸)$$

گیریم $F(x,y)=u$ و $F(x,z)=v$. آن‌گاه معادلهٔ (۸) به صورت $F(u,z)=F(x,v)$ در می‌آید. با مشتق‌گیری از این معادله به نوبت نسبت به x، y، و z، و با نوشتن $F_1(p,q)$ به جای $\partial F(p,q)/\partial p$ و $F_2(p,q)$ به جای $\partial F(p,q)/\partial q$، به دست می‌آوریم

$$F_1(u,z)\partial u/\partial x=F_1(x,v) \qquad (۱۹)$$

$$F_1(u,z)\partial u/\partial y=F_2(x,v)\partial v/\partial y \qquad (۲۰)$$

$$F_2(u,z)=F_2(x,v)\partial v/\partial z \qquad (۲۱)$$

با مشتق‌گیری از معادلهٔ (۲۰) به نوبت نسبت به x، y، و z، با نوشتن $F_{11}(p,q)$ به جای $\partial F_1(p,q)/\partial p$، و با نمایش دیگر مشتق‌های دوم به طرزی مشابه، به دست می‌آوریم

$$\begin{aligned}F_{11}(u,z)(\partial u/\partial x)(\partial u/\partial y) &+ F_1(u,z)\partial^2 u/\partial x\partial y\\ &= F_{12}(x,v)\partial v/\partial y\end{aligned} \qquad (۲۲)$$

$$\begin{aligned}&F_{11}(u,z)(\partial u/\partial y)^2 + F_1(u,z)\partial^2 u/\partial y^2\\ =&F_{22}(x,v)(\partial v/\partial y)^2 + F_2(x,v)\partial^2 v/\partial y^2\end{aligned} \qquad (۲۳)$$

$$\begin{aligned}F_{12}(u,z)\partial u/\partial y &= F_{22}(x,v)(\partial v/\partial y)(\partial v/\partial z)\\ &+ F_2(x,v)\partial^2 v/\partial y\partial z\end{aligned} \qquad (۲٤)$$

پس از مشتق‌گیری از معادلهٔ (۱۹) نسبت به z، یا از معادلهٔ (۲۱) نسبت به x، به دست می‌آوریم

$$F_{12}(u,z)\partial u/\partial x=F_{12}(x,v)\partial v/\partial z \qquad (۲۵)$$

در بین معاله‌های (۲۰)، (۲۲)، ... ، (۲۵) اکنون می‌توانیم تابع‌های u و v را به غیر از مشتق‌های آنها حذف کنیم. بدین سان، با حذف $F_{12}(u,z)$ و $F_{12}(x,v)$ بین معادله‌های (۲۲)، (۲٤)، و (۲۵)، معادلهٔ زیر را پیدا می‌کنیم

$$\begin{aligned}&[F_{11}(u,z)(\partial u/\partial y)^2-F_{22}(x,v)(\partial v/\partial y)^2](\partial u/\partial x)(\partial v/\partial z)\\ &=F_2(x,v)(\partial^2 v/\partial y\partial z)(\partial v/\partial y)(\partial u/\partial x)\\ &-F_1(u,z)(\partial^2 u/\partial x\partial y)(\partial u/\partial y)(\partial v/\partial z)\end{aligned}$$

پس از ترکیب این معادله با معادلهٔ (۲۳)، می‌توانیم $F_{11}(u,z)$ و $F_{22}(x,v)$ را با هم حذف کنیم و به دست آوریم

$$\begin{aligned}&F_1(u,z)(\partial v/\partial z)[(\partial^2 u/\partial y^2)(\partial u/\partial x)-(\partial^2 u/\partial x\partial y)(\partial u/\partial y)]\\ =&F_2(x,v)(\partial u/\partial x)[(\partial^2 v/\partial y^2)(\partial v/\partial z)-(\partial^2 v/\partial y\partial z)(\partial v/\partial y)]\end{aligned}$$

از راه ترکیب این معادله با معادلهٔ (۲۰)، می‌توانیم $F_1(u,z)$ و $F_2(x,v)$ را با هم حذف کرده به دست آوریم

$$\frac{\partial^2 u/\partial x\partial y}{\partial u/\partial x}-\frac{\partial^2 u/\partial y^2}{\partial u/\partial y}=\frac{\partial^2 v/\partial y\partial z}{\partial v/\partial z}-\frac{\partial^2 v/\partial y^2}{\partial v/\partial y}$$

این معادله را به صورت زیر می‌توان نوشت

$$\frac{\partial}{\partial y}ln(\frac{\partial u/\partial x}{\partial u/\partial y})=-\frac{\partial}{\partial y}ln(\frac{\partial v/\partial y}{\partial v/\partial z})$$

اینک $u=F(x,y)$ و $v=F(y,z)$، به طوری که

$$\frac{\partial u/\partial x}{\partial u/\partial y}=\frac{F_1(x,y)}{F_2(x,y)} \quad \text{و} \quad \frac{\partial v/\partial y}{\partial v/\partial z}=\frac{F_1(y,z)}{F_2(y,z)}$$

پس داریم

$$\frac{\partial}{\partial y}ln[\frac{F_1(x,y)}{F_2(x,y)}]=-\frac{\partial}{\partial y}ln[\frac{F_1(y,z)}{F_2(y,z)}]$$

چون x تنها در عضو سمت چپ و z تنها در عضو سمت راست این معادله ظاهر می‌شود، نتیجه می‌شود که هر عضو تابعی از تنها متغیر باقیمانده y است راحت‌تر خواهد بود که انتگرال این تابع را با $ln\Phi(y)$ نشان دهیم، به طوری که داشته باشیم

$$\frac{\partial}{\partial y}ln[\frac{F_1(x,y)}{F_2(x,y)}]=\frac{d}{dy}ln\Phi(y) \qquad (۲٦)$$

و

$$\frac{\partial}{\partial y}ln[\frac{F_1(y,z)}{F_2(y,z)}]=-\frac{d}{dy}ln\Phi(y) \qquad (۲۷)$$

با جای گردانی x، y، و z در معادلهٔ (۲۷)، به دست می‌آوریم

$$\frac{\partial}{\partial x}ln[\frac{F_1(x,y)}{F_2(x,y)}]=-\frac{d}{dx}ln\Phi(x) \qquad (۲۸)$$

پس از ضرب کردن معادلهٔ (۲۸) در dx و معادلهٔ (۲٦) در dy و جمع کردن آنها به دست می‌آوریم

$$\frac{\partial}{\partial x}ln[\frac{F_1(x,y)}{F_2(x,y)}]dx \quad + \quad \frac{\partial}{\partial y}ln[\frac{F_1(x,y)}{F_2(x,y)}]dy$$
$$= -d\ ln\Phi(x) \quad + \quad d\ ln\Phi(y)$$

چون عضو سمت چپ یک دیفرانسیل کامل است، می‌توانیم انتگرال بگیریم و بدین ترتیب به دست آوریم

$$F_1(x,y)/F_2(x,y) = h\Phi(y)/\Phi(x) \qquad (۲۹)$$

که در آن h ثابت انتگرال‌گیری است.

برای استفاده از این نتیجه، معادلهٔ (۲۰) را به معادلهٔ (۲۱) تقسیم کرده به دست می‌آوریم

$$\frac{F_1(u,z)}{F_2(u,z)}\frac{\partial u}{\partial y} = \frac{\partial v/\partial y}{\partial v/\partial z}$$

عضو سمت راست صرفاً برابر $F_1(y,z)/F_2(y,z)$ است و به کمک معادلهٔ (۲۹)، معادلهٔ بالا را می‌توان چنین نوشت

$$\frac{\Phi(z)}{\Phi(u)}\frac{\partial u}{\partial y} = \frac{\Phi(x)}{\Phi(y)}$$

در این معادله به جای u مقدار آن $F(x,y)$ را می‌گذاریم و داریم

$$\partial F(x,y)/\partial y = \Phi[F(x,y)]/\Phi(y) \qquad (۳۰)$$

به روال مشابه، از معادله‌های (۱۹) و (۲۰) به دست می‌آوریم

$$\partial F(y,z)/\partial y = \Phi[F(y,z)]/\Phi(y)$$

که وقتی x و y به جای y و z نوشته شوند، تبدیل می‌شود به

$$\partial F(x,y)/\partial x = \Phi[F(x,y)]/\Phi(x) \qquad (۳۱)$$

با ترکیب معادله‌های (۳۰) و (۳۱) برای به دست آوردن دیفرانسیل dF پیدا می‌کنیم

$$dF/\Phi(F) = dx/\Phi(x) + dy/\Phi(y)$$

(فهمیده می‌شود که متغیرها x و yاند.)

اگر $\int[dp/\Phi(p)]$ را با $ln\ f(p)$ نشان دهیم، پس از انتگرال‌گیری و نمایی کردن دو عضو این معادله،به دست می‌آوریم

$$Cf(F) = f(x)f(y)$$

که در آن C ثابت انتگرال‌گیری است. آن‌گاه این همان جواب معادلهٔ (۸) است.

جواب معادلهٔ

$$xS[S(y)/x] = yS[S(x)/y] \qquad (۱۵)$$

به روالی مشابه اما به طور سریعتر به دست می‌آید. این معادله و سه معادله‌ای راکه با مشتق‌گیری از آن نسبت به x، نسبت به y، و نسبت به x و y به دست می‌آیند، می‌توان به شرح زیر به دست آورد، اگر $S(y)/x$ را با u و $S(x)/y$ را با v نشان دهیم، داریم

$$xS(u) = yS(v) \qquad (۳۲)$$

$$uS'(u) - S(u) = -S'(v)S'(x) \qquad (۳۳)$$

$$S'(u)S'(y) = -vS'(v) + S(v) \qquad (۳٤)$$

$$uS''(u)S'(y)/x = vS''(v)S'(x)/y \qquad (۳۵)$$

پس از ضرب کردن معادلهٔ (۳۲) در معادلهٔ (۳۵)، x و y را به طور همزمان حذف می‌کنیم و به دست می‌آوریم

$$uS''(u)S(u)S'(y) = vS''(v)S(v)S'(x)$$

با این معادله همراه بامعادله‌های (۳۳) و (۳٤)، ممکن است که $S'(x)$ و $S'(y)$ را حذف کرد. نتیجه معادلهٔ زیر است

$$\frac{uS''(u)S(u)}{[uS'(u)-S(u)]S'(u)} = \frac{vS''(v)S(v)}{[vS'(v)-S(v)]S'(v)}$$

چون هر عضو معادلهٔ بالا تابعی مشابه از متغیری متفاوت است، این تابع باید برابر مقداری ثابت باشد. اگر این ثابت را k بنامیم، داریم

$$uS''(u)S(u) = k[uS'(u) - S(u)]S'(u)$$

این معادله را به صورت زیر می‌توان در آورد

$$dS'/S' = k(dS/S - du/u)$$

به کمک انتگرال‌گیری، به دست می‌آوریم

$$S' = A(S/u)^k$$

که در آن A ثابت انتگرال‌گیری است.

چون متغیرها جداشدنی‌اند، انتگرال‌گیری دیگر نتیجه می‌دهد

$$S^m = Au^m + B$$

که در آن m به جای $k-1$ نوشته شده، و B ثابت انتگرال‌گیری است. اکنون از راه جایگذاری می‌توان دریافت که اگر $B = A^2$ باشد این جواب به ازای مقدارهای دلخواه x و y در معادلهٔ (۱۵) صدق می‌کند. سرانجام، اگر قرار باشد جواب معادلهٔ (۱۵) در معادلهٔ $S[S(x)] = x$ نیز صدق کند، لازم است که $A = -1$ باشد. بدین ترتیب $[S(u)]^m + u^m = 1$ را به دست می‌آوریم.

مراجع

[1] Bergmann, G. (1941). The Logic of Probability. American Journal of Physics: **9 (5)**, 263-272.

[2] Boole, G. (1854). An Investigation of the Laws of Thought. Macmillan, London.

[3] Eddington, A. S. (1936). Relative Theory of Protons and Electrons. University Press, Cambridge.

[4] Hardy,G. H.(1920). Some Famous Problems of the Theory of Numbers and in Particular Waring's Problem. Clarendon Press, Oxford.

[5] Jefferys, H. (1939). Theory of Probability. Clarendon Press, Oxford.

[6] Kemble, E. C. (1942). Is the Frequency Theory of Probability Adequate for All Scientific Purposes?. American Journal of Physics, **10(1)**, 6-15.

[7] Keynes, J. M. (1929). A Treaties on Probability, Macmillan, London.

[8] Margenau, H. (1942). The Role of Definitions in Physical Science, with Remarks on the Frequency Definition of Probability. American Journal of Physics, **10(5)**, 224-232.

[9] McColl, H. (1880). The Calculus of Equivalent Statements (4th paper). Proceedings of London Mathematical Society, **11**, 113-121.

[10] Russell, B. (1972). Philosophy. Norton, New York.