آمار بیزی

Bayesian Statistics

آمار بیزی نظریه‌ی آماری منسجمی است که عدم قطعیت شناخت شناسیک[1] را با استفاده از زبان احتمال توصیف می‌کند. ، در این نظریه درجه‌های باور به وضعیت طبیعت اساس کار است. این درجه‌های باور با اعداد نامنفی بیان می‌شوند و مجموع آن‌ها دربار ۀ همۀ وضعیت‌های طبیعت باید برابر یک باشد. روش‌های آماری بیزی با باورهای پیشینی [2] قبل از مشاهده شروع می‌شوند. با بهره گیری از داده‌های آماری گردآوری شده از راه مشاهده یا آزمایش این باورها را روزآمد می‌کنند. نتیجۀ تحلیل بیزی باورهای پسینی[3] اند که می توانند به عنوان اساس تصمیم‌های استنباطی به‌کاروند.

این نظریه به نام توماس بیز کشیش پروتستان انگلیسی (۱۷۰۱–۱۷۶۱) نامیده شده است. کار بیزدر ۱۷۶۳ توسط دوستش ریچاردپرایس (۱۷۲۳–۱۷۹۱) انتشار یافت. هدف بیز آن بود که از داده‌های دو جمله‌ای متشکل از r پیروزی در بین n آزمایه، احتمال زیر ساختی پیروزی در هر آزمایه را بیابد.

سهم اساسی بیز ابتکار در به کارگیری توزیع احتمال برای نمایش عدم قطعیت دربارۀ پارامتر احتمال پیروزی (θ) بود. به خاطر نبود دانش کافی درباره‌ی θ این توزیع بیانگر عدم قطعیتی متفاوت از عدم قطعیت پیشامدهای آتی است. برای تمایز بین این دومفهوم، عدم قطعیت θ را با احتمال شناخت شناسیک و عدم قطعیت بنیادی[4] پیشامدهای تصادفی مانند آمدن شیریا خط درپرتاب سکّه را با احتمال بختیک[5] توصیف می کنند. عدم قطعیت شناخت شناسیک، بیانگر میزان آشنایی پیشینی آزمایشگر با پارامتر یا مقدارسرشتی یک پدیده است. درحالی که احتمال بختیک از غیرقابل‌پیشگویی بودن پیشامدهای آتی حکایت می‌کند. برای مثال، درپرتاب یک سکّه، یک مقدار سرشتی برای هرسکّه وجود دارد که موجب می‌شود شمار شیرها در n پرتاب کم یا زیاد شوند. اما درجۀ باور ما به این مقدار سرشتی عددی بین صفر و یک و باور به احتمال آمدن شیر در هر پرتاب است و عدم قطعیت شناخت شناسیک را تشکیل می‌دهد. مقدار سرشتی هرسکّه مقداری ثابت است اما نسبت به میزان آشنایی ما با ماهیت سکّه وطرز ساخت آن، درجۀ باورمان درباره‌ی آن مقدار سرشتی، می‌تواندهرعددی بین صفرویک باشد. درحالی‌که مفهوم عدم قطعیت بختیک یا احتمال آمدن شیر در هر پرتاب بیانگرآن است که اساساً به یقین نمی‌توان گفت درهربارپرتاب سکّه، کدام روی آن ظاهر خواهد شد و ربطی به باور ما ندارد. عدم قطعیت درباره کمیت‌های مجهول با توزیع‌های احتمال نمایش داده می شود.

مدل آماری به ازای هر θ توزیعی را برای متغیر تصادفی X مشخص می‌سازد که آن رابا $P(x|\theta)$ نشان می‌دهند. این توزیع، عدم قطعیت درباره‌ی داده x را پیش از معلوم شدن آن بر اثرمشاهده یا آزمایش بیان می‌کند به شرط آنکه مقدار پارامتر θ معلوم باشد. پارامتر θ هم توزیع خود $P(\theta)$ را دارد که بازتاب عدم قطعیت پژوهشگر درباره‌ی θ است .در واقع تحلیل‌گر $P(\theta)$ را به θ نسبت می‌دهد که جهل یا آ گاهی پیشینی اش را بیان کند. توزیع پسینی θ که پس از مشاهده دادۀ x مدنظر پژوهشگر است، طبق قضیۀ بیز به دست می‌آید. قضیۀ بیز قضیه‌ای برای احتمال شرطی است:

اگر دو کمیت تصادفی X و θ را داشته باشیم که عدم قطعیت درباره‌ی آنها به ترتیب با $P(x|\theta)$ و $P(\theta)$ توصیف شود، آن‌گاه:

$$P(\theta|x) = P(x|\theta)P(\theta) / \int P(x|\theta)P(\theta)\, d\theta$$

در مخرج کسر توزیع ناشرطی X را داریم که عدم قطعیت حقیقی درباره‌ی X را قبل از مشاهده داده‌ها توصیف می‌کند. اگر θ تنها چند مقدار متمایز را انتخاب کند، در مخرج به جای انتگرال از مجموع استفاده می‌شود.

قضیه بیز پایه‌ای برای استنباط بیزی درباره‌ی θ است. توزیع

پیشینی ($P(\theta)$) با درستنمایی ($P(x|\theta)$) ترکیب می‌شود و توزیع پسینی($P(\theta|x)$) را به دست می‌دهد.
در حالی که آمار بیزی نظریه‌ای غیر قابل مناقشه است، کاربست عملی آن مستلزم رعایت ملاحظاتی از موضوع‌های عملی از جمله انتخاب توزیع پیشینی، گزینش تابع درستنمایی، محاسبه و تلخیص توزیع پسینی در مسئله‌های چندبعدی است.
نظریه بیزی ازویژگی انسجام برخوردار است. نظریه‌های رقیب از این امتیاز بی‌بهره‌اند. از این رو، کاربرد احتمال به مثابه زبانی برای بیان عدم قطعیت ها ، گزینه‌ای اختیاری نیست و از اصولی ژرف‌تر دربارهٔ استدلال منطقی یا رفتار عقلانی ناشی می‌شود (دوفینتی ۱۹۷۴،۱۹۷۵).
در مباحث آمار بسامدگرا کراراً از استقلال متغیرهای تصادفی سخن به میان می‌آید. به ویژه در مبحث پیشگویی از پیشگویی مقدار متغیر مستقل X_{n+1}، پس از مشاهدهٔ x_1، x_2، . . . ، x_n سخن می‌رود. باید توجه داشت که این استقلال به شرط مقداری معین از پارامتر معنا دارد و در واقع استقلال شرطی است. وگرنه غیرمنطقی و ناممکن است برای پیشگویی مقدار یک متغیر از مقدارهایی استفاده شود که مستقل از آن هستند. این نکته اهمیت شرطی بودن مشاهدات نمونه‌ای و استفاده از آنها برای پیشگویی مقدار آینده را نشان می‌دهد که مستلزم متوسط گیری نسبت به عدم قطعیت شناخت شناسیک پارامتراست.
نکته‌ای اساسی که در استنباط‌های آماری قابل‌توجه پژوهشگران است موضوع انسجام است. آمار بسامدگرا مجموعه‌ای از روش‌هایی است که با ارتباطاتی ضعیف به هم پیوسته‌اند. از این رو آمار ورزان باید انتخاب کنند که برای حل مسئله خود از برآورد نقطه‌ای یا از آزمون فرض استفاده کنند. خاصیت اساسی‌ای که حتما لازم است روش‌های آماری داشته باشند این است که منسجم باشند. وضعیتی نامنسجم در آمار بسامدگرا برای توزیع نرمال در حالت دو متغیری و تک متغیری پیش می‌آید که گاهی آزمون‌های جداگانه برای میانگین های تک متغیرهٔ $M_1 = 0$ و $M_2 = 0$ رد می‌شوند اما آزمون دو متغیره‌ی $M_1 = M_2 = 0$ پذیرفته می‌شود. بنابراین آزمون بسامد گرا راهی نامنسجم برای تحلیل داده‌ها ست.
اصل‌های موضوع انسجام به نتیجه‌های زیر می‌انجامند که باید در روش‌های آماری رعایت شوند:

- عدم قطعیت‌ها با احتمال توصیف شوند.
- پیامدهای اقدامات و تصمیم‌ها با مفهوم مطلوبیت توصیف شوند.
- بهترین اقدام یا تصمیم آن باشد که مطلوبیت موردنظر را ماکسیمم می‌سازد.

در آمار بیزی باید از توزیع پیشینی استفاده کرد اما توزیع پیشینی چه می‌تواند باشد. پژوهشگر آزاد است هر پیشینی را که صلاح می‌داند برگزیند. اصولا انتخاب پیشین باید دانش پیشینی پژوهشگر دربارهٔ پدیدهٔ تحت بررسی را به درستی بازتاب دهد. اما پیشین به هر روشی که انتخاب شود تحلیل را در معرض انتقاد ذهنی بودن قرار می‌دهد. پاسخ بیزی گرایان به این انتقاد چندگونه است. برخی مثلا مانند برگر و بری(۱۹۸۸)، ذهنی بودن را سودمند می‌دانند. می‌گویند به این دلیل دانشمندان مختلف به نتیجه‌های متفاوت می‌رسند که اطلاعات پیشینی متفاوت دارند و بیان صریح پیشین (برخلاف قبول ضمنی برخی مفروضات درآمارکلاسیک) کاری پسندیده است. برخی دیگر از آمارشناسان بیزی با پیشنهاد توزیع‌های پیشینی که در برخی ملاک‌های عینی بودن صدق می‌کنند نیاز به تحلیل عینی بودن را تصدیق کرده برآنند که چنین پیشین‌هایی باید به عنوان پیشین‌های ازپیش تعیین شده برای مواردی که اطلاعات پیشینی ضعیف‌اند یا در مواردی که عدم توافق شدید درباره‌ی ماهیت توزیع پیشینی وجوددارد، به کارروند. در مواردی که داده‌های نمونه اطلاعات زیادی برای تمایز بین مقدارهای

مختلف پارامتر نداشته وتابع درستنمایی نسبتا پخش باشد ، توزیع پسینی نسبت به انتخاب پیشین خیلی حساس خواهد بود. در مسئله‌های دیگر که اطلاعات نمونه قابل توجه باشد، تابع درست‌نمایی کشیده خواهد بودو اثر پیشین را تحت‌الشعاع قرارخواهدداد. در این موارد حساسیت پسین نسبت به تغییرات پیشین اندک خواهد بود.

یکی از نقاط قوت آمار بیزی آسانی آن در پیشگویی[۶] است. متناظر با توزیع پسینی $P(\theta|x)$ ، توزیع پیشگوی[۷] هر کمیت y که از طریق توزیع نمونه‌گیری $P(y|\theta)$ به θ وابسته باشد به شرح زیر تعیین می شود:

$$P(y|x)=\int P(y|\theta)P(\theta|x)\,d\theta$$

به شرط آنکه x و y به شرط معلوم بودن θ مستقل شرطی باشند. این شرط به جز در مورد سری‌های زمانی ومدل‌های فضایی، معمولا برقرار است. در این دومورد از توزیع توام نمونه استفاده می شود.

نظریهٔ تصمیم بیزی

برای هدف‌های استنباط آماری، گزارشی کامل از توزیع پسینی کافی است اما ممکن است غیرعملی باشد به ویژه وقتی که پسین چندبعدی است. به جای آن معمولا خلاصه‌های پسینی مانند میانگین، واریانس وغیره گزارش می‌شوند. اگر تحلیل با هدف تصمیم گیری خاص انجام گیرد، ملاک‌های مطلوبیت یا تابع‌های زیان به کارمی‌روند تا در پرتوداده‌ها بهترین تصمیم اتخاذ شود. تابع زیان[۸] مقدار زیان ناشی از کار بست کنش خاص a را زمانی‌که مقدارواقعی پارامتر θ است نشان می‌دهد و به صورت $L(a,\theta)$ مشخص می‌شود. ضابطهٔ $L(a,\theta)$ به مسئلهٔ مورد بحث و نظر پژوهشگر در مورد زیان به بارآمده براثرکنش نادرست بستگی دارد.

به‌ازای هر پسین خاص، قاعدهٔ بیزی تصمیمی است که زیان مورد انتظار را نسبت به آن پسین مینیمم می‌سازد.اگر یک قاعدهٔ تصمیم گیری پذیرفتنی[۹] باشد (به این معنا که هیچ قاعده‌ای به ازای دست کم مقداری از پارامتر اکیدا بهتر از آن نباشد) می‌توان نشان داد که یک قاعده بیزی حاصل از نوعی پیشین سره و تابع مطلوبیت است (رابرت، ۲۰۰۷). بسیاری از خلاصه‌های آماری توزیع‌های پسینی که به لحاظ شهودی معقول‌اند نیز می‌توانند با قاعده‌های بیزی در چارچوب نظریهٔ تصمیم به دست آیند. از این رو برخی آمار بیزی را در چارچوب نظریهٔ تصمیم بیزی می‌پسندند، رابرت (۲۰۰۷). برای مثال، امیدریاضی پسینی پارامتر، قاعده‌ای بیزی برای وقتی است که تابع زیان درجه دوم $L(\theta,a)=(\theta-a)^2$ به کار می‌رود. بازه‌های باورمند[۱۰] مبتنی بر چندک‌های پسینی قاعده‌ای بیزی برای مسئله تصمیم دو بعدی‌اند (شرویش، ۱۹۹۵).

در تحلیل بیزی، محاسبهٔ مقدارمتوسط کمیت‌های تصادفی، به عنوان پایه‌ای برای استنباط آماری لازم است. این کمیت‌ها ممکن است شامل توزیع‌های پسینی چند متغیره با فرمول‌های پیچیده باشند که با قواعد انتگرال‌گیری معمولی نتوان از عهدهٔ آن‌ها برآمد. پیشرفت روش‌های محاسباتی با استفاده از رایانه‌های سریع تحولی درآمار بیزی ایجادکرده است که دیگر محدودیتی برای محاسبه وجود ندارد. تقریبا همه محاسبات را می‌توان به کمک روش‌های محاسباتی رایانه‌یار مونت‌کارلوی زنجیرمارکوفی[۱۱] (MCMC) انجام داد (رابرت و کاسه لا، ۲۰۰۴). نرم‌افزار WINBUGS امروزه ابزار تحلیل بیزی پژوهشگران در رشته‌های بسیار متنوع علمی شده است.

جفریز (۱۹۳۹) شیوه‌ای را برای آزمون فرضیه‌های علمی بدیل H_0 و H_1 ابداع کرده است. عامل بیزی[۱۲] که برپایه نسبت بخت‌های پسینی به پیشینی تحت هریک از فرضیه‌ها تعریف می‌شود ابزار این آزمون است. بخت، نسبت احتمال

رخداد یک پیشامد به احتمال عدم رخداد آن است. گیریم H_0 و H_1 به ترتیب با مجموعه های A و B از مقدارهای θ متناظر بوده وآن دو مجموعه مجزا از هم باشند. به ازای پیشین‌های مشخص شده برای θ تحت H_0 و H_1 که با $P_0(\theta)$ و $P_1(\theta)$ نمایانده می‌شوند، عامل بیزی عبارت‌است از

$$\int p(x|\theta)p(\theta)I(\theta \varepsilon A)d\theta / \int p(x|\theta)p(\theta)I(\theta \varepsilon B)$$

اگر H_0 و H_1 شامل تنها یک مقدار باشند، یعنی با آزمون فرض ساده $\{H_0 = \theta = \theta_0\}$ در مقابل فرض ساده $\{H_1 : \theta = \theta_1\}$ سروکارداشته باشیم، $B_{01}(x)$ به $P(x|\theta_0)/P(x|\theta_1)$ فرو می‌کاهد که همان نسبت درستنمایی آشنا و اساس روش آزمون فرض بسامدی نیمن – پیرسون را تشکیل می‌دهد. برای تصمیم گیری و انتخاب بین H_0 و H_1 جفریز(۱۹۳۹) آستانه هایی از قوت شواهد را له یا علیه فرض‌ها پیشنهاد کرده است.

کتاب‌شناسی

Berger, J. O. and Berry D. A. (1998). Statistical Analysis and the Illusion of Objectivity. *American Statistician*,***76***, pp. 59-73.

de Finetti Bruno (1974, 1975). *Theory of Probability*. London: Wiley.

Jeffreys, H. (1939). *Theory of Probability*. Oxford University Press, Oxford.

Price, Richard (1763). An Essay toward Solving a Problem in the Doctrines of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price. In a letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society*, **53**, pp. 370-418.

Robert, Christian, P. (2007). *The Bayesian Choice: Form Decision –Theoretic Foundations to Computational Implementation (2nd ed.)*, New York: Springer.

Robert, Christian, P. and Casella, G. (2004). *Monte Carlo Statistical Methods (2nd ed.)*. New York: Wiley.

Scherwish, M. J. (1995). *Theory of Statistics*. New York: Springer.

دانش‌واژه‌ها:

1. Epistemological uncertainty
2. Prior beliefs = a priori beliefs
3. Posterior beliefs = a posteriori beliefs
4. Essential uncertainty
5. Aleatory probability
6. Prediction
7. Predictive distribution
8. Loss function
9. Admissible
10. Credible intervals
12. Markov Chain Monte Carlo
13. Bayes factor

محمدرضا مشکانی

هیئت علمی دانشگاه شهید بهشتی