

Bayesian Statistics

آمار بیزی نظریه آماری منسجمی است که قطعیت نداشتن شناخت شناسیک^۱ 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 درباره همه وضعیت‌های طبیعت باید برابر یک باشد. روش‌های آماری بیزی با باورهای پیشینی^۲ قبل از مشاهده شروع می‌شوند. با بهره‌گیری از داده‌های آماری گردآوری شده از راه مشاهده یا آزمایش این باورها را روزآمد می‌کنند. نتیجه تحلیل بیزی باورهای پسینی^۳ اند که می‌توانند به عنوان اساس تصمیم‌های استنباطی به کار روند.

این نظریه به نام توماس بیز کشیش پروتستان انگلیسی (۱۷۰۱-۱۷۶۱) نامیده شده است. کار بیز در ۱۷۶۳ میلادی به وسیله دوستش ریچارد پرایس (۱۷۲۳-۱۷۹۱) انتشار یافت. هدف بیز آن بود که از داده‌های دوجمله‌ای متشکل از r پیروزی در بین n آزمایش، احتمال زیرساختی پیروزی در هر آزمایش را بیابد.

سهم اساسی بیز ابتکار در به کارگیری توزیع احتمال برای نمایش قطعیت نداشتن درباره پارامتر احتمال پیروزی (θ) بود. به خاطر نبود دانش کافی درباره θ این توزیع بیانگر قطعیت نداشتنی متفاوت از قطعیت نداشتن پیشامدهای آتی است. برای تمایز بین این دو مفهوم، قطعیت نداشتن θ را با احتمال شناخت شناسیک و قطعیت نداشتن بنیادی^۴ پیشامدهای تصادفی مانند آمدن شیر یا خط در پرتاب سکه را با احتمال بختیک^۵ توصیف می‌کنند. قطعیت نداشتن شناخت

شناسیک بیانگر میزان آشنایی پیشینی آزمایشگر با پارامتر یا مقدار سرشتی یک پدیده است. درحالی که احتمال بختیک از غیرقابل پیشگویی بودن پیشامدهای آتی حکایت می‌کند. برای مثال، در پرتاب یک سکه یک مقدار سرشتی برای هر سکه وجود دارد که موجب می‌شود شمار شیرها در n پرتاب کم یا زیاد شوند. اما درجه باور ما به این مقدار سرشتی عددی بین صفر و یک و باور به احتمال آمدن شیر در هر پرتاب است و قطعیت نداشتن شناخت شناسیک را تشکیل می‌دهد. مقدار سرشتی هر سکه مقداری ثابت است اما در مقایسه با میزان آشنایی ما با ماهیت سکه و طرز ساخت آن، درجه باورمان درباره آن مقدار سرشتی، می‌تواند هر عددی بین صفر و یک باشد. درحالی که مفهوم قطعیت نداشتن بختیک یا احتمال آمدن شیر در هر پرتاب بیانگر آن است که اساساً به یقین نمی‌توان گفت در هر بار پرتاب سکه کدام روی آن ظاهر خواهد شد و ربطی به باور ما ندارد. قطعیت نداشتن درباره کمیت‌های مجهول با توزیع‌های احتمال نمایش داده می‌شود.

مدل آماری به ازای هر θ توزیعی را برای متغیر تصادفی X مشخص می‌سازد که آن را با $P(x|\theta)$ نشان می‌دهند. این توزیع، قطعیت نداشتن درباره داده X را پیش از معلوم شدن آن بر اثر مشاهده یا آزمایش بیان می‌کند به شرط آنکه مقدار پارامتر θ معلوم باشد. پارامتر θ هم توزیع خود $P(\theta)$ را دارد که بازتاب قطعیت نداشتن پژوهشگر درباره θ است. درواقع تحلیل گر $P(\theta)$ را به θ نسبت می‌دهد که جهل یا آگاهی پیشینی‌اش را بیان کند. توزیع پسینی θ که پس از مشاهده داده X مدنظر پژوهشگر است، طبق قضیه بیز به دست می‌آید. قضیه بیز قضیه‌ای برای احتمال شرطی است:

اگر دو کمیت تصادفی X و θ را داشته باشیم که قطعیت

نداشتن درباره آنها به ترتیب با $P(x|\theta)$ و $P(\theta)$ توصیف شود، آن گاه:

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

در مخرج کسر توزیع ناشروطی X را داریم که قطعیت نداشتن حقیقی درباره 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 درباره θ است. توزیع پیشینی $(P(\theta))$ با درست‌نمایی $(P(x|\theta))$ ترکیب می‌شود و توزیع پسینی $(P(\theta|x))$ را به دست می‌دهد. درحالی‌که آمار بیزی نظریه‌ای غیرقابل مناقشه است، کاربرست عملی آن مستلزم رعایت ملاحظات از موضوع‌های عملی از جمله انتخاب توزیع پیشینی، گزینش تابع درست‌نمایی، محاسبه و تلخیص توزیع پسینی در مسئله‌های چندبعدی است.

نظریه بیزی از ویژگی انسجام برخوردار است. نظریه‌های رقیب از این امتیاز بی‌بهره‌اند. ازاین‌رو، کاربرد احتمال به مثابه زبانی برای بیان قطعیت نداشتن‌ها، گزینه‌ای اختیاری نیست و از اصولی ژرف‌تر درباره استدلال منطقی یا رفتار عقلانی ناشی می‌شود (دوفیتی، ۱۹۷۵ و ۱۹۷۴).

در مباحث آمار بسامدگرا به‌طور مکرر از استقلال متغیرهای تصادفی سخن به میان می‌آید. به‌ویژه در مبحث پیشگویی از پیشگویی مقدار متغیر مستقل X_{n+1} ، پس از مشاهده $x_1, x_2, \dots, x_n$ سخن می‌رود. باید توجه داشت که این استقلال به شرط مقداری معین از پارامتر معنا دارد و درواقع 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 متوسط‌گیری نسبت به قطعیت نداشتن شناخت شناسیک پارامتر است.

نکته‌ای اساسی که در استنباط‌های آماری قابل توجه پژوهشگران است موضوع انسجام است. آمار بسامدگرا مجموعه‌ای از روش‌هایی است که با ارتباطاتی ضعیف به هم پیوسته‌اند. ازاین‌رو، آمارورزان باید انتخاب کنند که برای حل مسئله خود از برآورد نقطه‌ای یا از آزمون فرض استفاده کنند. خاصیت اساسی‌ای که حتماً لازم است روش‌های آماری داشته باشند این است که منسجم باشند. وضعیتی نامنسجم در آمار بسامدگرا برای توزیع نرمال در حالت دومتغیری و تک‌متغیری پیش می‌آید که گاهی آزمون‌های جداگانه برای میانگین‌های تک‌متغیره $M_1=0$ و $M_2=0$ رد می‌شوند اما آزمون دومتغیره $M_1=M_2=0$ پذیرفته می‌شود. بنابراین، آزمون بسامدگرا راهی نامنسجم برای تحلیل داده‌هاست. اصل‌های موضوع انسجام به نتیجه‌های زیر می‌انجامند که باید در روش‌های آماری رعایت شوند:

- قطعیت نداشتن‌ها 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. اصولاً انتخاب

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 سری های زمانی و مدل های فضایی، معمولاً برقرار است. در این دومورد از توزیع توام نمونه استفاده می شود.

نظریه تصمیم بیزی

برای هدف های استنباط آماری، گزارشی کامل از توزیع پسینی کافی است اما ممکن است غیرعملی باشد به ویژه وقتی که پسین چندبعدی است. به جای آن معمولاً خلاصه های پسینی مانند میانگین، واریانس و غیره گزارش می شوند. اگر تحلیل با هدف تصمیم گیری خاص انجام گیرد، ملاک های مطلوبیت یا تابع های زیان به کار می روند تا در پرتو داده ها بهترین تصمیم اتخاذ شود. تابع زیان^۸ مقدار زیان ناشی از کار بست کنش خاص a را زمانی که مقدار واقعی پارامتر θ است نشان می دهد و به صورت $L(a, \theta)$ مشخص می شود. ضابطه $L(a, \theta)$ به مسئله مورد بحث و نظر پژوهشگر در مورد زیان به بارآمده بر اثر کنش نادرست بستگی دارد.

به ازای هر پسین خاص، قاعده بیزی تصمیمی است که زیان مورد انتظار را نسبت به آن پسین مینیم می سازد. اگر یک قاعده تصمیم گیری پذیرفتنی^۹ باشد (به این معنا که هیچ قاعده ای به ازای دست کم مقداری از پارامتر اکیداً بهتر از آن نباشد) می توان نشان داد که یک قاعده بیزی حاصل از نوعی پیشین سره و تابع مطلوبیت است (رابرت، ۲۰۰۷). بسیاری از

پیشین باید دانش پیشینی پژوهشگر درباره پدیده تحت بررسی را به درستی بازتاب دهد. اما پیشین به هر روشی که انتخاب شود تحلیل را در معرض انتقاد ذهنی بودن قرار می دهد. پاسخ بیزی گرایان به این انتقاد چندگونه است. برخی مثلاً مانند برگر و بری (۱۹۸۸)، ذهنی بودن را سودمند می دانند. می گویند به این دلیل دانشمندان متعدد به نتیجه های متفاوت می رسند که اطلاعات پیشینی متفاوت دارند و بیان صریح پیشین (برخلاف قبول ضمنی برخی مفروضات در آمار کلاسیک) کاری پسندیده است. برخی دیگر از آمارشناسان بیزی با پیشنهاد توزیع های پیشینی که در برخی ملاک های عینی بودن صدق می کنند نیاز به تحلیل عینی بودن را تصدیق کرده، برآنند که چنین پیشین هایی باید به عنوان پیشین های از پیش تعیین شده برای مواردی که اطلاعات پیشینی ضعیف اند یا در مواردی که اختلاف شدید درباره ماهیت توزیع پیشینی وجود دارد، به کار روند. در مواردی که داده های نمونه اطلاعات زیادی برای تمایز بین مقدارهای مختلف پارامتر نداشته و تابع درستنمایی نسبتاً پخش باشد، توزیع پسینی نسبت به انتخاب پیشین خیلی حساس خواهد بود. در مسئله های دیگر که اطلاعات نمونه قابل توجه باشد، تابع درستنمایی کشیده خواهد بود و اثر پیشین را تحت الشعاع قرار خواهد داد. 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 نمونه گیری $P(y|\theta)$ به θ

مجموعه‌های A و B از مقدارهای θ متناظر بوده و آن دو مجموعه مجزا از هم باشند. به‌ازای پیشین‌های مشخص‌شده برای θ تحت H_0 و H_1 که با $P_0(\theta)$ و $P_1(\theta)$ نمایانده می‌شوند، عامل بیزی عبارت است از:

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

اگر H_0 و H_1 شامل تنها یک مقدار باشند، یعنی با آزمون فرض ساده $\{H_1: \theta = \theta_1\}$ در مقابل فرض ساده $\{H_0: \theta = \theta_0\}$ سروکار داشته باشیم، $B_{0,1}(x)$ 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 تصمیم‌گیری و انتخاب بین H_0 و H_1 جفریز (۱۹۳۹) آستانه‌هایی از قوت شواهد را برای 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

خلاصه‌های آماری توزیع‌های پسینی که به‌لحاظ شهودی معقول‌اند نیز می‌توانند با قاعده‌های بیزی در چارچوب نظریه تصمیم به‌دست آیند. ازاین‌رو برخی آمار بیزی را در چارچوب نظریه تصمیم بیزی می‌پسندند (رابرت، ۲۰۰۷). برای مثال، امید ریاضی پسینی پارامتر، قاعده‌ای بیزی برای وقتی است که تابع زیان درجه دوم $L(\theta, a) = (\theta - a)^2$ به‌کار می‌رود. بازه‌های باورمند^{۱۱} مبتنی بر چندک‌های پسینی قاعده‌ای بیزی برای مسئله تصمیم دوبعدی‌اند (شرویش، ۱۹۹۵).

در تحلیل بیزی، محاسبه مقدار متوسط کمیت‌های تصادفی، به‌عنوان پایه‌ای برای استنباط آماری لازم است. این کمیت‌ها ممکن است شامل توزیع‌های پسینی چندمتغیره با فرمول‌های پیچیده باشند که با قواعد انتگرال‌گیری معمولی نتوان از عهده آنها برآمد. پیشرفت روش‌های محاسباتی با استفاده از رایانه‌های سریع تحولی در آمار بیزی ایجاد کرده است که دیگر محدودیتی برای محاسبه وجود ندارد. تقریباً همه محاسبات را می‌توان به‌کمک روش‌های محاسباتی رایانه‌ار 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 رخ ندادن آن است. گیریم H_0 و H_1 به‌ترتیب با

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



دانشگاه اقتصاد