

X1-089 (2013/01/28)

سنجش احتمال

mamwad@mailaps.org

محمد خرمی

تخمین احتمال و عدم قطعیت آن، و نیز اثر آزمایش و پیش-فرض بر این تخمینها بررسی میشود.

0 مقدمه

یک تعریف غیر-رسمی احتمال این است: یک رویداد را n بار می‌آزماییم. اگر m بار رخ داد، احتمال رخ-دادن آن رویداد (m/n) است. دقیقتر-شده‌ی این تعریف آن است که این نتیجه وقت‌ی درست است که تعداد آزمایشها زیاد باشد. این سئالها پیش می‌آیند.

- «تعداد آزمایشها زیاد باشد» یعنی چه؟ یک N هست که اگر تعداد آزمایشها بیشتر از N باشد تعداد آزمایشها زیاد (کافی) است؟
با فرض این که چنین N ی هست،
- اگر تعداد آزمایشها از N بیشتر باشد احتمال رویداد دقیق معلوم است؟
- اگر تعداد آزمایشها از N کمتر باشد در باره‌ی احتمال رویداد هیچ نمیدانیم؟

• اگر از پیش چیزی در باره ی احتمال آن رویداد بدانیم، آن داده چه طر وارد سنجش میشود؟

لابد جواب سه سئال اول منفی است. یک سکه (که از سالم-بودن ش مطمئن نیستیم) اگر ده بار شیر بیاید هم ممکن است بار یازدهم خط بیاید. آن ده بار که شیر آمده نشان نمیدهد احتمال شیر-آمدن ش یک است. اگر فقط دُ بار آزمایش کنیم و هر دُ بار نتیجه شیر باشد، حدس میزنیم احتمال شیر-آمدن بیشتر باشد، ولی ن آن قدر که از ده بار آزمایش نتیجه میگیریم.

در باره ی سئال آخر، اگر بدانیم سکه سالم است (یعنی بدانیم احتمال شیر-یا-خط-آمدن ش 0.5 است)، آن گاه نتیجه ی هیچ آزمایشی این داده (پیشفرض) را عوض نمیکند: اگر یک سکه ی سالم صد بار پشت سر هم شیر بیاید هم، احتمال شیر-آمدن ش 0.5 است ن 1. اما اگر به سالم-بودن سکه شک داشته باشیم، هر چه تعداد موارد آزمایش بیشتر شود (به شرطی که نتیجه ی آزمایش با پیشفرض متفاوت باشد) شک مان تقویت میشود.

دنبال مدل ی برای سنجش احتمالیم که ویژه گیها ی بالا را داشته باشد: با افزایش تعداد آزمایشها مقداری که برای احتمال میدهد قطعتر باشد، و با افزایش تعداد آزمایشها اثر پیشفرض کمتر شود، مگر از اول احتمال را با قطعیت بدانیم.

1 قضیه ی بیز

احتمال رخ-دادن \mathbb{A} را با $P(\mathbb{A})$ نشان میدهم. احتمال رخ-دادن \mathbb{A} و \mathbb{B} را با $P(\mathbb{A} \cap \mathbb{B})$ ، احتمال رخ-دادن \mathbb{A} یا \mathbb{B} را با $P(\mathbb{A} \cup \mathbb{B})$ ، و احتمال رخ-دادن \mathbb{B} به شرط \mathbb{A} را با $P(\mathbb{B}|\mathbb{A})$ نشان میدهم. داریم

$$P(\mathbb{A} \cup \mathbb{B}) = P(\mathbb{A}) + P(\mathbb{B}) - P(\mathbb{B} \cap \mathbb{A}). \quad (1)$$

$$P(\mathbb{B}|\mathbb{A}) = \frac{P(\mathbb{B} \cap \mathbb{A})}{P(\mathbb{A})}. \quad (2)$$

مجموعه ها ی \mathbb{E}_i با این ویژه گیها را در نظر بگیرید.

$$\mathbb{E}_i \cap \mathbb{E}_j = \emptyset, \quad i \neq j. \quad (3)$$

$$\bigcup_i \mathbb{E}_i = \mathbb{M}, \quad (4)$$

که \mathbb{M} مجموعه ی مرجع است (که احتمال ش یک است). داریم

$$\begin{aligned} P(\mathbb{A}) &= P(\mathbb{A} \cap \mathbb{M}), \\ &= P[\mathbb{A} \cap (\cup_i \mathbb{E}_i)], \\ &= P[\cup_i (\mathbb{A} \cap \mathbb{E}_i)], \\ &= \sum_i P(\mathbb{A} \cap \mathbb{E}_i), \end{aligned} \quad (5)$$

که در برابری ی آخر (1) به کار رفته، و این که به خاطر (3)،

$$(\mathbb{A} \cap \mathbb{E}_i) \cap (\mathbb{A} \cap \mathbb{E}_j) = \emptyset, \quad i \neq j. \quad (6)$$

داریم

$$\begin{aligned} P(\mathbb{E}_i | \mathbb{A}) &= \frac{P(\mathbb{E}_i \cap \mathbb{A})}{P(\mathbb{A})}, \\ &= \frac{P(\mathbb{E}_i \cap \mathbb{A})}{\sum_j P(\mathbb{E}_j \cap \mathbb{A})}, \end{aligned} \quad (7)$$

که نتیجه میدهد

$$P(\mathbb{E}_i | \mathbb{A}) = \frac{P(\mathbb{E}_i) P(\mathbb{A} | \mathbb{E}_i)}{\sum_j P(\mathbb{E}_j) P(\mathbb{A} | \mathbb{E}_j)}. \quad (8)$$

به این رابطه قضیه ی بیز [1] میگویند. اینها را میشود در مثلن فصل 4 از [2] یافت.

وقت ی مجموعه ی مرجع پیوسته است و بر آن انتگرالگیری تعریف شده، چگالی ی احتمال را تابع ی تعریف میکنیم که انتگرال ش احتمال است. چگالی ی احتمال را با D نشان میدهم. در این صورت،

$$P(\mathbb{S}) =: \int_{\mathbb{S}} (dV)(\xi) D(\xi), \quad (9)$$

که dV اندازه ی انتگرالگیری است. متغیر تصادفی ی X نگاهت ی از مجموعه ی مرجع است. چگالی ی احتمال برا ی X را با D_X نشان میدهم:

$$\begin{aligned} \int_{\mathbb{U}} (dV_X)(x) D_X(x) &:= P[X^{-1}(\mathbb{U})], \\ &= \int_{X^{-1}(\mathbb{U})} (dV)(\xi) D(\xi). \end{aligned} \quad (10)$$

dV_X اندازه ی انتگرالگیری در مجموعه ی مقادیرها ی X ، یعنی در $X(\mathbb{M})$ ، است. به این ترتیب، اگر Y یک متغیر تصادفی باشد چنان که (X, Y) یک-به-یک باشد و

$$(dV)(\xi) = (dV_X)[X(\xi)] (dV_Y)[X(\xi), Y(\xi)], \quad (11)$$

آنگاه،

$$D_X(x) = \int_{X^{-1}(x)} dV_Y[x, Y(\xi)] D(\xi). \quad (12)$$

به هم ین ترتیب،

$$\begin{aligned} \int_{\mathbb{U}} (dV_X)(x) D_{X \cap \mathbb{A}}(x) &:= P\{[X^{-1}(\mathbb{U})] \cap \mathbb{A}\}, \\ &= \int_{[X^{-1}(\mathbb{U})] \cap \mathbb{A}} (dV)(\xi) D(\xi), \end{aligned} \quad (13)$$

که با تعریف $\Theta_{\mathbb{A}}$ با

$$\Theta_{\mathbb{A}}(\xi) := \begin{cases} 1, & \xi \in \mathbb{A} \\ 0, & \xi \notin \mathbb{A} \end{cases}, \quad (14)$$

نتیجه میدهد

$$D_{X \cap \mathbb{A}}(x) = \int_{[X^{-1}(x)] \cap \mathbb{A}} dV_Y[x, Y(\xi)] \Theta(\xi) D(\xi). \quad (15)$$

سرانجام،

$$\begin{aligned} P(\mathbb{A}) &= \int_{\mathbb{A}} (dV)(\xi) D(\xi), \\ &= \int_{\mathbb{M}} (dV)(\xi) \Theta(\xi) D(\xi), \\ &= \int_{\mathbb{M}} (dV_X)[X(\xi)] (dV_Y)[X(\xi), Y(\xi)] \Theta(\xi) D(\xi), \\ &= \int (dV_X)(x) D_{X \cap \mathbb{A}}(x), \end{aligned} \quad (16)$$

$$D_{X|\mathbb{A}} = \frac{D_{X \cap \mathbb{A}}}{P(\mathbb{A})}, \quad (17)$$

$$P(\mathbb{A}|X) = \frac{D_{X \cap \mathbb{A}}}{D_X}. \quad (18)$$

به این ترتیب،

$$D_{X|\mathbb{A}}(x) = \frac{D_X(x) [P(\mathbb{A}|X)](x)}{\int (dV_X)(x') D_X(x') [P(\mathbb{A}|X)](x')}. \quad (19)$$

2 احتمال یک رویداد، بر اساس آزمایش

رویداد \mathbb{A} را n بار می‌آزماییم. m بار رخ میدهد. خاسته احتمال رخ-دادن آن رویداد است. این احتمال را با متغیر تصادفی X نشان میدهیم. این که طی n آزمایش m بار رویداد \mathbb{A} رخ دهد - نظر

$$D_{X|\mathbb{A}_m|n}(x) = \frac{D_X(x) [P(\mathbb{A}_m|n|X)](x)}{\int (dV_X)(x') D_X(x') [P(\mathbb{A}_m|n|X)](x')}. \quad (20)$$

رخ بدهد را با \mathbb{A}_m نشان میدهیم. رابطه ی (19) میشود
 این احتمال $[P(\mathbb{A}_m|n|X)](x)$ این است که با n بار آزمایش m بار رویداد \mathbb{A} رخ دهد، به شرطی که احتمال رخ-دادن رویداد \mathbb{A} x باشد. اگر آزمایشها ناپسته باشند، این احتمال از تزیع \mathbb{A} به دست می‌آید:

$$[P(\mathbb{A}_m|n|X)](x) = \binom{n}{m} x^m (1-x)^{n-m}, \quad (21)$$

(مثلن فصل 3 از [2]). به این ترتیب،

$$(dV_X)(x) D_{X|\mathbb{A}_m|n}(x) = \frac{(dV_X)(x) D_X(x) x^m (1-x)^{n-m}}{\int dx' D_X(x') x'^m (1-x')^{n-m}}, \quad (22)$$

که با تعریف D_X و $D_{X|\mathbb{A}_m|n}$ با

$$dx D_X(x) := (dV_X)(x) D_X(x),$$

$$dx D_{X|\mathbb{A}_m|n}(x) := (dV_X)(x) D_{X|\mathbb{A}_m|n}(x), \quad (23)$$

میشود

$$D_{X|\mathbb{A}_m|n}(x) = \frac{D_X(x) x^m (1-x)^{n-m}}{\int_0^1 dx' D_X(x') x'^m (1-x')^{n-m}}, \quad (24)$$

$D_X(x)$ چگالی ی احتمال این است که احتمال رخ-دادن رویداد \mathbb{A} x باشد. x باشد، به این معنی که $dx D_X(x)$ احتمال این است که احتمال رخ-دادن رویداد \mathbb{A} x باشد. $D_{X|\mathbb{A}_m|n}(x)$ چگالی ی احتمال این است که طی یک آزمایش بین x و $x + dx$ رخ دهد.

احتمال رخ-دادن رویداد مُرد- نظر طی یک آزمایش x باشد، به شرط این که قبلاً طی n بار آزمایش، m بار آن رویداد رخ داده باشد.

با چگالی-ی-احتمال $D_{X|\mathbb{A}_{m|n}}$ میشود میانگین $f(X)$ را حساب کرد:

$$\langle f(X) \rangle_{\mathbb{A}_{m|n}} = \int_0^1 dx D_{X|\mathbb{A}_{m|n}}(x) f(x). \quad (25)$$

از جمله،

$$\langle X \rangle_{\mathbb{A}_{m|n}} = \int_0^1 dx D_{X|\mathbb{A}_{m|n}}(x) x, \quad (26)$$

$$[\text{Var}(X)]_{\mathbb{A}_{m|n}} = \int_0^1 dx D_{X|\mathbb{A}_{m|n}}(x) (x - \langle X \rangle_{\mathbb{A}_{m|n}})^2. \quad (27)$$

اما برای به-دست-آوردن $D_{X|\mathbb{A}_{m|n}}$ باید D_X را بدانیم، یعنی باید چگالی ی احتمال پیش از انجام آزمایش را بدانیم.

3 سنجش احتمال یک رویداد، بدون پیش-فرض

اگر هیچ پیش-فرض ی نداشته باشیم (یعنی مثلاً در مُرد پرتاب سکه ندانیم احتمال شیر-آمدن 0.5 یا نزدیک آن است)، یک انتخاب برای D_X این است که آن را در دامنه ی تعریف ش ثابت بگیریم:

$$D_X^{\text{uni}}(x) = 1, \quad 0 \leq x \leq 1. \quad (28)$$

به این ترتیب، با استفاده از

$$\int_0^1 dx x^m (1-x)^{n-m} = \frac{m! (n-m)!}{(n+1)!}, \quad (29)$$

نتیجه میشود

$$D_{X|\mathbb{A}_{m|n}}^{\text{uni}}(x) = \frac{(n+1)!}{m! (n-m)!} x^m (1-x)^{n-m}. \quad (30)$$

بیشینه ی این چگالی در x_0 رخ میدهد، که

$$x_0 = \frac{m}{n}. \quad (31)$$

همچنین،

$$\begin{aligned} \langle X^k \rangle_{\mathbb{A}_{m|n}}^{\text{uni}} &= \frac{(n+1)!}{m!(n-m)!} \int_0^1 dx x^{m+k} (1-x)^{n-m}, \\ &= \frac{(n+1)!}{m!(n-m)!} \frac{(m+k)!(n-m)!}{(n+k+1)!}, \\ &= \frac{(m+k)!(n+1)!}{m!(n+k+1)!}, \end{aligned} \quad (32)$$

که از جمله نتیجه میدهد

$$\langle X \rangle_{\mathbb{A}_{m|n}}^{\text{uni}} = \frac{m+1}{n+2}, \quad (33)$$

$$[\text{Var}(X)]_{\mathbb{A}_{m|n}}^{\text{uni}} = \frac{(m+1)(n-m+1)}{(n+2)^2(n+3)}. \quad (34)$$

دیده میشود مقدار چشمداشتی احتمال، لزوماً با x_0 (جا بی که چگالی بیشینه میشود) برابر نیست. البته با افزایش تعداد آزمایشها جا بی که چگالی بیشینه میشود و مقدار چشمداشتی احتمال به هم نزدیک میشوند:

$$\langle X \rangle_{\mathbb{A}_{m|n}}^{\text{uni}} - x_0 = \frac{n-2m}{n(n+2)}. \quad (35)$$

با افزایش تعداد آزمایشها، وردایی هم کم میشود. به این ترتیب، با افزایش تعداد آزمایشها اطلاعات مان در باره احتمال رخ-دادن رویداد بیشتر میشود، و مقداری که برای احتمال رخ-دادن رویداد تخمین میزنیم به هم ان مقداری که میگویند، یعنی به (m/n) ، میگراید.

4 سنجش احتمال یک رویداد، با پیش-فرض

وجود پیش-فرض، یعنی پیش از آزمایش چیزی در باره X میدانیم. مثلاً بعید است احتمال شیر-آمدن یک سکه خیل با 0.5 فرق کند. این اطلاعات را میشود در D_X وارد کرد. یک راه این است که این اطلاعات را هم-ارز با تعدادی آزمایش بگیریم. یعنی پیش-فرض را هم-ارز با این بگیریم که ν آزمایش انجام شده، و μ بار رویداد مُرد-نظر رخ داده. این یعنی $D_X^{\mathbb{A}_{\mu|\nu}}$ بگیریم، که

$$D_X^{\mathbb{A}_{\mu|\nu}}(x) = \frac{(\nu+1)!}{\mu!(\nu-\mu)!} x^\mu (1-x)^{\nu-\mu}. \quad (36)$$

اگر بخواهیم مقدار چشمداشتی X با این چگالی ξ باشد،

$$\xi = \frac{\mu+1}{\nu+2}. \quad (37)$$

این یک معادله برای μ متغیر ν و μ است. معادله دیگری لابد از این جا می آید که چه قدر به این که احتمال x است مطمئن یم. از رابطه ν و μ با تعداد آزمایشها دیده میشود هر چه تعداد آزمایشها بیشتر باشد و ردایی کمتر است، یعنی اطمینان مان بیشتر است. پس به عنوان یک معیار از مقدار اطمینان میشود ν را داد.

میشود به صحیح بودن ν و μ مقید نبود. البته μ و ν باید چنان باشند که $D_X^{\Delta\mu|\nu}$ انتگرال-پذیر باشد. این یعنی باید μ و $(\nu - \mu)$ بزرگتر از (-1) باشند. داریم

$$\mu = (\nu + 2)\xi - 1, \quad (38)$$

$$\nu - \mu = \nu + 1 - (\nu + 2)\xi. \quad (39)$$

به این ترتیب، اگر ξ صفر یا یک نباشد، شرط انتگرال-پذیری ν میشود

$$\nu > -2. \quad (40)$$

در حالتها بی که x_0 صفر یا یک است، در واقع احتمال را میدانیم: اگر یک متغیر تصادفی فقط بین صفر و یک مقدار بگیرد، و مقدار چشمداشتی ν صفر یا یک باشد، آنگاه چگالی ν احتمال برای آن متغیر دلتا ν دیرک [3] است. البته حالتها $\xi = 0$ و $\xi = 1$ را میشود با $\nu \rightarrow \infty$ و به ترتیب $(\mu/\nu) \rightarrow 1$ و $(\mu/\nu) \rightarrow 0$ هم به دست آورد.

سرانجام، دیده میشود

$$D_X^{\text{uni}} = D_X^{\Delta 0|0}. \quad (41)$$

این یعنی حالت بدون پیش-فرض، هم ارز با آن است که تعداد مثر آزمایشها ν متناظر با پیش-فرض صفر است، چنان که انتظار میرفت.

با گذاشتن $D_X^{\Delta\mu|\nu}$ به جای D_X در (24)، نتیجه میشود

$$D_{X|\Delta m|\nu}^{\Delta\mu|\nu}(x) = \frac{(n + \nu + 1)!}{(m + \mu)!(n + \nu - m - \mu)!} x^{m+\mu} (1-x)^{n+\nu-m-\mu}. \quad (42)$$

به این ترتیب، اثر پیش-فرض فقط آن است که n و m به، به ترتیب، $(n + \nu)$ و $(m + \mu)$ تبدیل

میشوند. از جمله،

$$\langle X \rangle_{\mathbb{A}_m | n}^{\mathbb{A}_{\mu | \nu}} = \frac{m + \mu + 1}{n + \nu + 2}, \quad (43)$$

$$[\text{Var}(X)]_{\mathbb{A}_m | n}^{\mathbb{A}_{\mu | \nu}} = \frac{(m + \mu + 1)(n + \nu - m - \mu + 1)}{(n + \nu + 2)^2 (n + \nu + 3)}. \quad (44)$$

رُشن است که با این مدل، هر چه ν بزرگتر باشد پیش-فرض مهمتر میشود، و هر چه n بزرگتر باشد آزمایشها مهمتر میشوند.

5 پانوشتها

- [1] Bayes
- [2] Morris H. deGroot & Mark J. Schervish; "probability and statistics" 4th edition (Addison-Wesley, 2002)
- [3] Dirac