



فرایند تصادفی احتمال و توزیع ها

محسن هوشمند
دانشکده تکنولوژی اطلاعات و علم رایانه
دانشگاه تحصیلات تکمیلی علوم پایه زنجان

استنتاج آماری

کمک در تصمیم بر آنچه بدان اعتقاد یابیم

جستجو بر یافتن تحقیق

مدل

▪ توصیف ریاضی

مدل آماری

▪ معمولا مدل احتمالی

مدل دارای پارامتر است

مثال

- احتمال باران وابسته به مثلا سطح دریا
- باران خروجی
- سطح دریا: ورودی

رابطه دقیق بین ورودی و خروجی با مقدار دیگری تدوین می شود که نمایشگر این است که ورودی چقدر بر خروجی تأثیر می گذارد

- پارامتر

تدوین مدل (معادله بندی مدل): نمایشگر اینکه ورودی بر خروجی تأثیر گذار است.

- پارامترها مقدار دقیق تأثیر را معین می کنند

مدل دارای پارامتر است

مثال

- شیر آمدن سکه
- سکه کج
- اختلاف بین دو سطح x
- $P(\text{ش}|x)$
- درجهٔ اعتماد
- هایپرپارامتر

سخن کوتاه

مدل ریاضی احتمالِ پیشامدهای خاصی که رخ می‌دهند.

مدل ریاضی دارای پارامترهایی است.

مقدار پارامترها مشخص‌کنندهٔ احتمال دقیق بدست‌آمده از مدل‌هاست

اعتقادات به مقدار پارامترها جریان می‌دهد.

شاید بعضی از مقادیر را نسبت به مقادیر دیگر بیشتر مستعد درستی بدانیم

صورت اعتقاد ما دربارهٔ مقدار پارامترها خود می‌تواند با مدل ریاضی دیگری بیان شود [هایپر]-پارامترها

اهداف استنتاج

تخمین مقدار پارامتر

پیش‌بینی مقدار داده

- کمک مدل‌ها به پیش‌بینی اثربخشی واکنش انفلوانزا هنگام واکنش عمومی
- کمک مدل‌ها به پیش‌بینی مسیر طوفان
- کمک مدل‌ها به پیش‌بینی شیر یا خط آمدن پرتاب بعدی سکه
- در استنتاج بیز، میانگین وزنی اعتقاد است.

مقایسه مدل (انتخاب مدل)

احتمال

روش‌های آماری مطالعه دقیقِ عدم قطعیت‌ها

نیاز به فهم احتمال

مجموعه متناهی پیشامدهای ممکن

- پرتاب سکه؟
- چقدر محتمل است؟
- خط چقدر محتمل است؟
- هر دو آمدن چقدر محتمل است؟
- محتمل بودن هر برآمد، مجموعه‌ای از برآمدها را در ذهن داریم
- هر دو آمدن جز پیشامدهای ممکن نیست

پیشامد

هنگام پرسش دربارهٔ محتمل شدن پیشامدی مجموعه‌ای از پیشامدها را در ذهن داریم، شامل تمامی پیشامدهای ممکن و پیشامدهای دو به دو مستقل از یکدیگر

- فضای نمونه

احتمال

مطالعه عدم قطعیت‌ها

نسبتی از زمان اتفاق افتادن پیشامد

درجه اعتقاد درباره پیشامد

عدم قطعیت در

- داده
- مدل یادگیر
- پیش‌بینی‌های حاصل از مدل

مفاهیم فضای احتمال

- فضای نمونه
- پیشامدها
- احتمال هر رخداد

احتمال

شیر آمدن

▪ احتمال شیر آمدن پارامتر θ

▪ سکه سالم $\theta = 0.5$

▪ درجه اعتقاد $P(\theta)$

▪ $P(\theta = 0.5) = 0.99$

▪ احتمال:

▪ شیر و خط

▪ فضای نمونه دارند

▪ درجه اعتقاد

▪ مقدار پیوسته $\theta = 0.001, 0.002, \dots, 0.01, 0.02, \dots, 0.1, 0.2, \dots, 0.9, 1.0$

چرا سکه

▪ چقدر سکه پرتاب می کنیم؟

▪ مشابه زندگی - احتمال زنده ماندن یکسال پس از جراحی

▪ احتمال اثر جانبی سردرد ناشی از دارو

▪ انتخابات یا رفراندوم

احتمال

نیاز به متغیر تصادفی جهت کمی‌سازی عدم قطعیت

- تابع واسط بین خروجی آزمایش تصادفی به مجموعه‌ای از ویژگی‌های مورد توجه

توزیع احتمال

- مرتبط با متغیر تصادفی
- تابع اندازه‌گیرنده احتمال خروجی خاص
- جزو اساسی
- مدل‌های احتمالاتی
- مدل‌های گرافیکی
- انتخاب مدل

ایجاد فضای احتمال

هدف نظریه احتمال

- تعریف ساختی ریاضی برای توضیح خروجی‌های تصادفی آزمایش‌ها
- مثال
 - پرتاب سکه
 - عدم امکان تشخیص خروجی
 - با انجام بسیاری آزمایش، مشاهده نظم در میانگین خروجی
 - پرتاب تاس
- استدلال خودکار
- تعمیم استدلال منطقی

روش‌های ترکیبیاتی

اعمال چند مرحله‌ای

- دو مرحله‌ای
- بیش از دو مرحله

جایگشت

- n شی متمایز
- چینش r شی از n شی متمایز
- n شی متمایز روی دایره
- n شی دارای k دسته متمایز

ترکیب (آرایش) (زیرمجموعه)

- انتخاب r شی از n شی متمایز
- تعداد افراهای n شی متمایز به r زیرمجموعه

ضرائب دو جمله‌ای

$$(x + y)^n = \sum_{r=0}^n \binom{n}{r} x^{n-r} y^r$$

$$\binom{n}{r} = \binom{n}{n-r}$$

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

$$\sum_{r=0}^n \binom{m}{r} \binom{n}{k-r} = \binom{m+n}{k}$$

$$n! \cong \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

معادله استرلینگ

تفسیرهای احتمال

کلاسیک (احتمال برابر)

بسامدی

بیزی

اصل آغاز

فضای نمونه Ω

آزمایش

- هر فرایند مشاهده یا اندازه‌گیری

برآمد

- نتایج حاصل از آزمایش

فضای نمونه

- مجموعه‌ء همه برآمدهای ممکن آزمایش
- فضای نمونه پرتاب سکه، اعداد صحیح فرد و مثبت، فضای نمونه‌ای که از پرتاب جفت تاس، یکی سرخ و دیگری سبز
- گسسته (متناهی و یا نامتناهی شمارا) در مقابل پیوسته (نقاط فضای نمونه تشکیل پیوستار)

پیشامد

- زیرمجموعه‌هائی از فضای نمونه
- نادرستی عکس مطلب

فضای پیشامد

- معمولا مجموعه توانی Ω

مثال

طول عمر مفید وسیله الکترونیکی
▪ فضای نمونه

وسیله الکترونیکی قبل از ششمین سال از کار بیفتد
▪ پیشامد

مثال

طول عمر مفید وسیله الکترونیکی
▪ فضای نمونه

وسیله الکترونیکی قبل از ششمین سال از کار بیفتد
▪ پیشامد

$$\Omega = \{t: t \geq 0\}$$

$$F = \{t: 0 \leq t \leq 6\}$$

تابع احتمال و اصول کولمولگروف

تابعی را که به هر پیشامد، عددی در بازه ۰ تا ۱ نسبت دهد و صادق در سه اصل زیر

نمایش احتمال پیشامد A با $P(A)$
▪ $P(A)$

- اصل اول: احتمال هر پیشامد نامنفی $P(A) \geq 0$
- اصل دوم: احتمال جمع پیشامدهای مجزا (ناسازگار) برابر با جمع احتمالهای تک تک آنها
- اصل سوم: احتمال فضای نمونه برابر با ۱

مثال - پرتاب تاس - فضای نمونه، پیشامد اعداد زوج؛ پیشامد اعداد کوچکتر از ۵

مثال

احتمال حداقل یک شیر آمدن دو بار پرتاب سکه متعادل

مثال

احتمال پیشامد وقوع عدد بزرگتر از ۳ برای تاسی که احتمال وقوع هر عدد فرد دو برابر احتمال وقوع هر عدد زوج است

ویژگی‌ها

A' متمم پیشامد A

$$p(A') = 1 - P(A)$$

ϕ مجموعه تهی

$$p(\phi) = 0$$

$$A \subseteq B \Rightarrow P(A) \leq P(B)$$

$$0 \leq P(A) \leq 1$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B \cup C)$$

$$= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

ویژگی‌ها

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A \cap B) \leq \min(P(A), P(B))$$

$$P(\phi) = 0$$

$$A \subseteq B \Rightarrow P(A) \leq P(B)$$

$$0 \leq P(A) \leq 1$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

احتمال شرطی

احتمال شرطی پیشامد A به شرط وقوع پیشامد B

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

به سخن دیگر، احتمال شرطی پیشامد A است بعد از مشاهده پیشامد B

$$p(A \cap B) = p(B)p(A|B) = p(A)p(B|A)$$

احتمال شرطی

نمودار ون

مثال

	عرضه خدمات مناسب	عرضه خدمات بد
سابقه شغلی کمتر از ۱۰ سال	۱۶	۴
سابقه شغلی بیشتر از ۱۰ سال	۱۰	۲۰

خ: احتمال مرکز مناسب چقدر است؟

$$P(\text{خ}) = \frac{26}{50} = 0.52$$

ک: در صورت انتخاب مرکز با کمتر از ۱۰ سال سابقه، احتمال مناسب بودن خدمات چقدر است؟

$$p(\text{خ}|\text{ک}) = \frac{p(\text{خ} \cap \text{ک})}{p(\text{ک})} = \frac{\frac{16}{50}}{\frac{20}{50}} = \frac{16}{20} = 0.8$$

قاعده زنجیری

پیشامدهای S_1, \dots, S_k و $P(S_i) > 0$

برای $k = 2$

$$P(S_1 \cap S_2) = P(S_1)P(S_2|S_1)$$

در واقع قاعده زنجیری مشتق از اعمال چندباره احتمال شرطی است.

برای $k = 3$

$$P(S_1 \cap S_2 \cap S_3) = P(S_1)P(S_2|S_1)P(S_3|S_1 \cap S_2)$$

برای k

$$P(S_1 \cap S_2 \cap \dots \cap S_k) = P(S_1)P(S_2|S_1)P(S_3|S_1 \cap S_2) \dots P(S_k|S_1 \cap S_2 \cap \dots \cap S_{k-1})$$

مثال

اگر از ۲۴۰ لامپ تلویزیونی، ۱۵ لامپ معیوب باشد، دو لامپ را بدون جایگذاری به ترتیب برداریم.
احتمال اینکه هر دو معیوب باشد چقدر است؟

خ: احتمال معیوب بودن اولین لامپ

ک: احتمال معیوب بودن دومین لامپ

$$p(x \cap k) = p(x)p(k|x) = \frac{15}{240} \frac{14}{239}$$

مثال

۲۰ فیوز در بسته داریم که ۵ مورد آنها معیوب است.

خ: احتمال معیوب بودن اولین لامپ

ک: احتمال معیوب بودن دومین لامپ

ل: احتمال معیوب بودن سومین لامپ

احتمال اینکه هر سه معیوب باشد چقدر است؟

$$p(x \cap k \cap l) = p(x)p(k|x)p(l|x \cap k) = \frac{5}{20} \frac{4}{19} \frac{3}{18}$$

استقلال

دو پیشامد مستقل خوانده می شوند اگر

$$P(A \cap B) = P(A)P(B)$$

یا

$$P(A | B) = P(A)$$

شهود دو پیشامد مستقل

▪ عدم تاثیر مشاهده B بر مشاهده A

مثال

۲۰ فیوز در بسته داریم که ۵ مورد آنها معیوب است.

خ: احتمال معیوب بودن اولین لامپ

ک: احتمال معیوب بودن دومین لامپ

ل: احتمال معیوب بودن سومین لامپ

اگر فیوز را برگردانیم. احتمال اینکه هر سه معیوب باشد چقدر است؟

$$p(x) = \frac{5}{20}$$

$$p(k|x) = \frac{5}{19}$$

$$p(l|x \cap k) = \frac{5}{18}$$

$$p(x \cap k \cap l) = p(x)p(k|x)p(l|x \cap k) = \left(\frac{5}{20}\right)^3$$

مثال

سکه‌ای را سه بار پرتاب کنیم. خ: ۱ ش: 0، $\Omega = \{000, 001, 010, \dots, 111\}$

احتمال پیشامد شیر در دو پرتاب اول α ، $p(\alpha) = \frac{1}{4}$

احتمال پیشامد خط در پرتاب سوم β ، $p(\beta) = \frac{1}{2}$

احتمال پیشامد دقیقا دو خط در سه پرتاب γ ، $p(\gamma) = \frac{3}{8}$

پیشامدهای α و β مستقل‌اند؟ $p(\alpha\beta) = \frac{1}{8} = p(\alpha)p(\beta)$

پیشامدهای γ و β مستقل‌اند؟ $p(\gamma \cap \beta) = \frac{1}{4}$ ، $p(\gamma)p(\beta) = \frac{3}{16}$

قضية بيز

مثال

ساخت بزرگراهی به علت عدم اعتبار ممکن است به تاخیر بیفتد. احتمال فراهم نشدن اعتبار ۰,۶ است. احتمال انجام سروقت بزرگراه به شرط وجود اعتبار ۰,۸۵ و احتمال ان به شرط عدم اعتبار ۰,۳۵ است.

$$p(\beta) = 0.6$$

$$p(\alpha|\beta) = 0.35$$

$$p(\alpha|\beta') = 0.85$$

$$\begin{aligned} p(\alpha) &= p((\alpha \cap \beta) \cup (\alpha \cap \beta')) \\ &= p(\alpha \cap \beta) + p(\alpha \cap \beta') \\ &= p(\beta)p(\alpha|\beta) + p(\beta')p(\alpha|\beta') \\ &= 0.6 \times 0.35 + 0.4 \times 0.85 = 0.55 \end{aligned}$$

قاعده احتمال کل (قاعده حذف)

پیشامدهای $\beta_1, \beta_2, \dots, \beta_k$ افرازی از فضای نمونه Ω و $p(\beta_i) \neq 0$

$$p(\alpha) = \sum_{i=1}^k p(\beta_i)p(\alpha|\beta_i)$$

مثال

در استان زنجان دودزائی اتومبیل‌ها بیش از حد استاندارد است. اتومبیل دودزا در آزمون با احتمال ۰,۹۹ مردود و اتومبیل سالم در آزمون با احتمال ۰,۱۷ مردود می‌شود. احتمال اینکه اتومبیل در آزمون رد شود واقعا دودزا باشد چقدر است؟

$$p(\beta) = ۰.۲۵ \text{ اتومبیل دودزا}$$

$$p(\beta') = ۰.۷۵ \text{ اتومبیل سالم}$$

رد شدن α

$$p(\alpha|\beta) = ۰.۹۹$$

$$p(\alpha|\beta') = ۰.۱۷$$

$$p(\beta \cap \alpha) = p(\alpha)p(\beta|\alpha) = p(\beta)p(\alpha|\beta) \Rightarrow p(\beta|\alpha) = \frac{p(\beta)p(\alpha|\beta)}{p(\alpha)}$$

$$p(\beta|\alpha) = \frac{p(\beta)p(\alpha|\beta)}{p(\beta)p(\alpha|\beta) + p(\beta')p(\alpha|\beta')}$$

قاعدهٔ بیز

$$p(\beta|\alpha) = \frac{p(\beta)p(\alpha|\beta)}{p(\beta)p(\alpha|\beta) + p(\beta')p(\alpha|\beta')}$$

تعمیم

پیشامدهای $\beta_1, \beta_2, \dots, \beta_k$ افرازی از فضای نمونه Ω و $p(\beta_i) \neq 0$

$$p(\beta_i|\alpha) = \frac{p(\beta_i)p(\alpha|\beta_i)}{\sum_{j=1}^k p(\beta_j)p(\alpha|\beta_j)}$$

متغیر تصادفی

توجه به جنبه‌ای خاص

▪ فرض ریختن جفت تاس و صرفاً توجه به مجموعه رقم‌ها

▪ (1,1)

▪ (2,1)

▪ ...

عدم استفاده مستقیم از فضای احتمال در مسائل عموماً و خاصه هوش مصنوعی

فضای هدف

$$X: \Omega \rightarrow \mathcal{T}$$

▪ متغیر تصادفی

▪ غلط مصطلح و سوءتفاهم

▪ نه تصادفی نه متغیر

▪ بلکه تابع

متغیرهای تصادفی RANDOM VARIABLES مت

مثال - پرتاب ۱۰ سکه و به دنبال اینکه تعداد شیرهایی که از این تعداد پرتاب به دست می‌آید.
▪ اعضای فضای نمونه دنباله‌های ده‌تایی از شیر و خط هستند

در عمل به دنبال این «نیستیم» که احتمال دنباله‌ای خاص چقدر است.
▪ به دنبال توابع مقدار-حقیقی از خروجی‌ها هستیم.

▪ مثلاً مجموع دو عددی که در ریختن جفت تاس، یا تعداد شیرهایی که در ۱۰ پرتاب تاس به دست می‌آید

مت:

$$X: \Omega \rightarrow \mathbb{R}$$

نمایش با $X(\omega)$ یا X

$X = x$ یعنی مقدار $x \in R$ را به مت X اختصاص داده‌ایم

متغیرهای تصادفی - ادامه

اگر Ω فضای نمونه با اندازه احتمال باشد و X تابع حقیقی-مقدار روی اعضای فضای نمونه باشد-

مثال

سکه‌ای چهار با پرتاب می‌شود. مت تعداد کل شیرها

خ: ۱ ش: ۰، $\Omega = \{0000, 0001, 0010, \dots, 1111\}$

$$P(X = 3) = \frac{4}{16}$$

فضای نمونه	احتمال	$X=x$
0000	$\frac{1}{16}$	۴
0001	$\frac{1}{16}$	۳
0010	$\frac{1}{16}$	۳
0011	$\frac{1}{16}$	۲
0100	$\frac{1}{16}$	۳
0101	$\frac{1}{16}$	۲
0110	$\frac{1}{16}$	۲
0111	$\frac{1}{16}$	۱
1000	$\frac{1}{16}$	۳
1001	$\frac{1}{16}$	۲
1010	$\frac{1}{16}$	۲
1011	$\frac{1}{16}$	۱
1100	$\frac{1}{16}$	۲
1101	$\frac{1}{16}$	۱
1111	$\frac{1}{16}$	۰

متغیرهای تصادفی - ادامه

اگر Ω فضای نمونه با اندازه احتمال باشید و X تابع حقیقی-مقدار روی اعضای فضای نمونه باشد

مثال متغیر تصادفی گسسته: تعداد شیر در پرتاب ۱۰ سکه

▪ مقادیر محدودی را می‌پذیرد

$$P(X = k) = P(\{\omega : X(\omega) = k\})$$

متغیر تصادفی پیوسته: زمانی نیمه عمر ماده رادیواکتیو؛ طول عمر مفید وسیله‌الکترونیکی

▪ مقداری نامتناهی را می‌پذیرد

▪ a و b به طوری که $a < b$

$$P(a \leq X \leq b) = P(\{\omega : a \leq X(\omega) \leq b\})$$

x	$P(X = x)$
۲	$\frac{۱}{۳۶}$
۳	$\frac{۲}{۳۶}$
۴	$\frac{۳}{۳۶}$
۵	$\frac{۴}{۳۶}$
۶	$\frac{۵}{۳۶}$
۷	$\frac{۶}{۳۶}$
۸	$\frac{۵}{۳۶}$
۹	$\frac{۴}{۳۶}$
۱۰	$\frac{۳}{۳۶}$
۱۱	$\frac{۲}{۳۶}$
۱۲	$\frac{۱}{۳۶}$

مثال - توزیع احتمال

فضای نمونه پرتاب جفت تاس هر عضو فضا $\frac{۱}{۳۶}$
متغیر تصادفی حاصل جمع دو عدد تاس‌ها

تدوین معادله

$$P(X = x) = \frac{۶ - |x - ۷|}{۳۶}, x = ۲, ۳, \dots, ۱۲$$

توزیع‌های احتمال

فهرستی از تمامی خروجی‌های ممکن و احتمال‌های متناظر هر یک

- توزیع احتمال پرتاب سکه-

- دو خروجی شیر و خط

- احتمال‌های متناظر آنها θ و $1 - \theta$

- توزیع احتمال مقدار کالری مصرفی روزانه انسان

- ۲۰۰۰.۰

- ۱۸۹۸.۳

- ۲۴۴۷.۹

- ...

- توزیع‌های گسسته

- فضای نمونه حاوی خروجی گسسته و احتمال مجزای هر برآمد

- پرتاب سکه، ریختن تاس

- تقسیم فضای پیوسته

- توزیع‌های پیوسته

- احتمال نقطه‌ای صفر

- احتمال بازه‌ای ممکن

- چگالی احتمال

توزیع‌ها احتمال

بر اساس اصول کولمولگروف

قضیه: توابعی را می‌توان به مثابه توزیع احتمالی متغیر تصادفی گسسته بکار برد که مقادیر $f(x)$ اگر و فقط اگر مشمول شرایط زیر باشند

$$f(x) \geq 0$$
$$\sum_x f(x) = 1$$

مثال

تدوین معادله برای توزیع احتمال تعداد کل شیرهای چهار پرتاب سکه

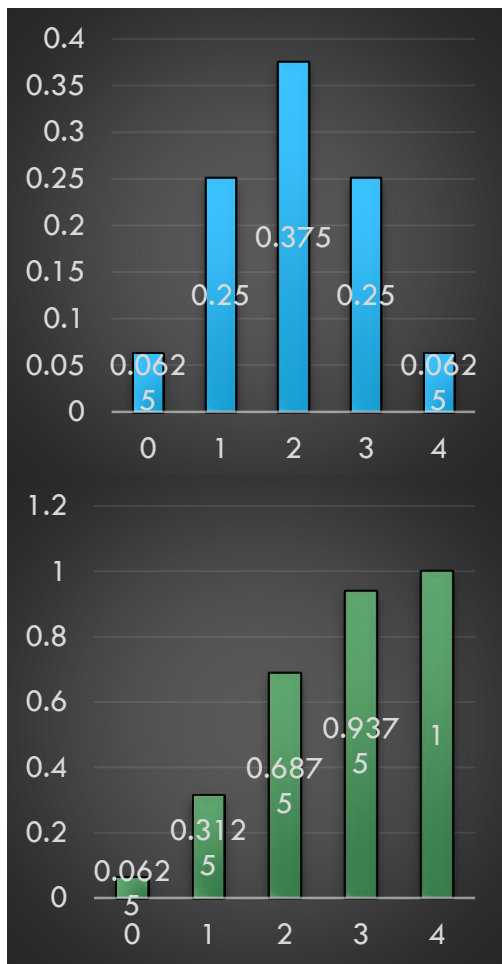
$$\begin{aligned}P(X = 0) &= \frac{1}{16} \\P(X = 1) &= \frac{4}{16} \\P(X = 2) &= \frac{6}{16} \\P(X = 3) &= \frac{4}{16} \\P(X = 4) &= \frac{1}{16}\end{aligned}$$

تدوین معادله

$$P(X = x) = \frac{\binom{4}{x}}{16}, x = 0, 1, 2, 3, 4$$

مثال - نمودار میله‌ای

$$P(X = 0) = \frac{1}{16}$$
$$P(X = 1) = \frac{4}{16}$$
$$P(X = 2) = \frac{6}{16}$$
$$P(X = 3) = \frac{4}{16}$$
$$P(X = 4) = \frac{1}{16}$$



تدوین معادله برای توزیع احتمال تعداد کل شیرهای چهار پرتاب سکه

مثال

$$P(x) \geq 0$$
$$\sum_x P(x) = 1$$

بررسی کنید تابع زیر تابع توزیع است.

$$f(x) = \frac{x + 2}{25}, x = 1, 2, 3, 4, 5$$

$$f(x) \geq 0$$
$$\sum_x f(x) = 1$$

توابع توزیع تجمیعی - مثال

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad -\infty \leq x \leq +\infty$$

$$F(+\infty) = 1$$

مثال

توزیع احتمال این تابع تجمیعی چقدر است؟

$$F(x) = \begin{cases} 0, & x < 2 \\ \frac{1}{36}, & 2 \leq x < 3 \\ \frac{2}{36}, & 3 \leq x < 4 \\ \frac{3}{36}, & 4 \leq x < 5 \\ \frac{4}{36}, & \vdots \\ \frac{5}{36}, & 4 \leq x < 5 \\ \frac{6}{36}, & x \geq 12 \end{cases}$$

$$P(2) = \frac{1}{36}$$

$$P(3) = \frac{2}{36} - \frac{1}{36} = \frac{1}{36}$$

$$P(4) = \frac{3}{36} - \frac{2}{36} = \frac{1}{36}$$

\vdots

$$P(12) = 1 - \frac{35}{36} = \frac{1}{36}$$

x	$f(X = x)$
2	$\frac{1}{36}$
3	$\frac{2}{36}$
4	$\frac{3}{36}$
5	$\frac{4}{36}$
6	$\frac{5}{36}$
7	$\frac{6}{36}$
8	$\frac{5}{36}$
9	$\frac{4}{36}$
10	$\frac{3}{36}$
11	$\frac{2}{36}$
12	$\frac{1}{36}$

توابع توزیع تجمیعی - ادامه

ویژگی‌ها

$$0 \leq F(X) \leq 1$$

$$\lim_{x \rightarrow -\infty} F(X) = 0$$

$$\lim_{x \rightarrow +\infty} F(X) = 1$$

$$x \leq y \implies F(x) \leq F(y)$$

توابع توزیع تجمیعی - ادامه

قضیه: اگر متغیر تصادفی متشکل از مقادیر $x_1 < x_2 < \dots < x_n$ باشد، آن گاه $P(x_1) = F(x_1)$ و

$$P(x_i) = F(x_i) - F(x_{i-1}), i = 2, 3, \dots, n$$

مثال

$$P(X = 0) = \frac{1}{16}$$

$$P(X = 1) = \frac{4}{16}$$

$$P(X = 2) = \frac{6}{16}$$

$$P(X = 3) = \frac{4}{16}$$

$$P(X = 4) = \frac{1}{16}$$

تعداد کل شیرهای چهار پرتاب سکه

$$F(0) = P(0) = \frac{1}{16}$$

$$F(1) = P(0) + P(1) = \frac{5}{16}$$

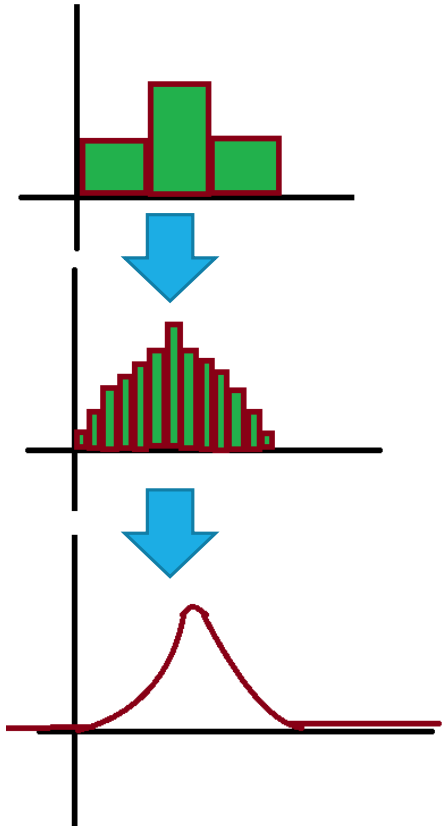
$$F(2) = P(0) + P(1) + P(2) = \frac{11}{16}$$

$$F(3) = P(0) + P(1) + P(2) + P(3) = \frac{15}{16}$$

$$F(4) = P(0) + P(1) + P(2) + P(3) + P(4) = 1$$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{16}, & 0 \leq x < 1 \\ \frac{5}{16}, & 1 \leq x < 2 \\ \frac{11}{16}, & 2 \leq x < 3 \\ \frac{15}{16}, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

تابع چگالی احتمال PROBABILITY DENSITY FUNCTION «تچا» یا «PDF»



در حالت گسسته: $P(x_i) = F(x_i) - F(x_{i-1})$

اگر مت پیوسته باشد، ت^۳ در همه جا مشتق پذیر است

«تچا» مشتق «ت^۳» است (شبیه تفاضل در حالت گسسته)

$$f_X(x) = \frac{dF_X(x)}{dx}$$
$$P(x \leq X \leq x + \delta x) = f_X(x)\delta x$$

توجه: مقدار تچا مربوط به یک نقطه، احتمال آن نقطه نیست

$$f_X(x) \neq P(X = x)$$

تابع چگالی احتمال - ویژگی‌ها

$$f_X(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$\int_{x \in A} f_X(x) dx = P(X \in A)$$

مثال

$f(x) = \begin{cases} ke^{-3x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$ در چه صورتی چگالی احتمال است؟

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$\int_{-\infty}^{+\infty} ke^{-3x} dx = \int_0^{+\infty} ke^{-3x} dx = \lim_{t \rightarrow \infty} \left. \frac{ke^{-3x}}{-3} \right|_0^t = \frac{k}{3} = 1 \Rightarrow k = 3$$

$$F(x) = P(X \leq x) = \int_{-\infty}^x f_X(x) dx$$

▪ تابع توزیع تجمیعی

تابع چگالی احتمال -

قضیه: اگر $f(x)$ و $F(x)$ توابع توزیع احتمال و تابع توزیع X به ازای x باشند، آن گاه به ازای هر دو عدد حقیقی a و b با شرط $a \leq b$:

$$P(a \leq x \leq b) = F(b) - F(a)$$

$$f_X(x) = \frac{dF_X(x)}{dx}$$

و این مشتق موجود است

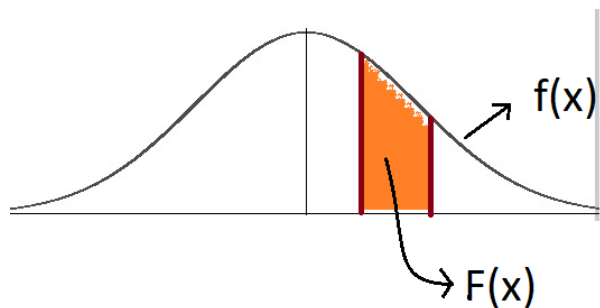
مثال

تابع چگالی احتمال را بدست آورید و رسم کنید.

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

مثال

فرض کنید ۱۰۰۰ باتری را امتحان کردیم و نمودار آنها شبیه زیر است



احتمال برابر سطح زیر منحنی در حد فاصل مقدار کم و بیش بازه

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx$$

تابع توزیع احتمال

گسسته

- تفاضل
- تابع جرم احتمال

$$P(x) \geq 0$$
$$\sum_x P(x) = 1$$

$$0 \leq P(x) \leq 1$$

- چرا؟

پیوسته

- مشتق
- تابع چگالی احتمال

$$f_X(x) \geq 0$$
$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$f_X(x) \leq 1$$

- خیر

دو متغیر تصادفی

توزیع‌های توام و حاشیه‌ای JOINT AND MARGINAL DISTRIBUTIONS

تاکنون صرفاً یک متغیر تصادفی

ممکن است که بیش از یک متغیر تصادفی درگیر باشد
▪ مثال

X و Y دو متغیر تصادفی و $F_X(x)$ و $F_Y(y)$ توزیع متناظر آنها

نیاز به توزیعی قوی‌تری برای فهم کار آنها با یکدیگر
▪ تابع توزیع جرمی توام X و Y

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

توزیع‌های توام - مثال

دو قرص به تصادف از بسته‌ای محتوای سه قرص اسپرین و ۲ قرص خواب‌آور و ۴ قرص سرماخوردگی انتخاب می‌کنیم.

X تعداد قرص اسپرین، Y تعداد قرص خواب‌آور

X	Y
0	0
0	۱
۱	0
۱	۱
0	۲
۲	0

$$P(X = 1, Y = 0) = \frac{\binom{3}{1} \binom{2}{0} \binom{4}{1}}{\binom{9}{2}} = \frac{1}{3}$$

$$P(X = 1, Y = 1) = \frac{\binom{3}{1} \binom{2}{1} \binom{4}{0}}{\binom{9}{2}} = \frac{1}{6}$$

$$P(X = x, Y = y) = \frac{\binom{3}{x} \binom{2}{y} \binom{4}{2-x-y}}{\binom{9}{2}}, x, y \in \{0, 1, 2\}, 0 \leq x + y \leq 2$$

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$
$Y = 1$	$\frac{2}{9}$	$\frac{1}{6}$	
$Y = 2$	$\frac{1}{36}$		

ویژگی‌ها

$$f(x, y) \geq 0$$
$$\sum_x \sum_y f(x, y) = 1$$

مثال ۱، ۲ و ۳ $f(x, y) = kxy, \quad x, y = 1, 2, 3$

توزیع تجمیعی توام X و Y

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{s \leq x} \sum_{t \leq y} f(s, t)$$

توزیع‌های توام - مثال

دو قرص به تصادف از بسته‌ای محتوای سه قرص آسپرین و ۲ قرص خواب‌آور و ۴ قرص سرماخوردگی انتخاب می‌کنیم.

X تعداد قرص آسپرین، Y تعداد قرص خواب‌آور

$F(1,1)$ ؟

X	Y
0	0
0	۱
۱	0
۱	۱
0	۲
۲	0

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$
$Y = 1$	$\frac{2}{9}$	$\frac{1}{6}$	
$Y = 2$	$\frac{1}{36}$		

حالت پیوسته

$f(x,y)$ تابع چگالی احتمال توام اگر و فقط اگر

$$P[(X,Y) \in A] = \int_A \int f(x,y) dx dy$$

ویژگی‌ها

$$\begin{aligned} f(x,y) &\geq 0 \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) dx dy &= 1 \end{aligned}$$

تابع توزیع توام

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^x f(s, t) ds dt$$

چگالی توام:

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$$

مثال

چگالی توام X و Y

$$f(x, y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1 \\ 0 & \text{سایر جاها} \end{cases}$$

$$F(x, y) = \begin{cases} 0 & , x \leq 0, y \leq 0 \\ \frac{1}{2}xy(x + y), & 0 < x < 1, 0 < y < 1 \\ \frac{1}{2}y(1 + y), & x \geq 1, 0 < y < 1 \\ \frac{1}{2}x(x + 1), & 0 < x < 1, y \geq 1 \\ 1, & x \geq 1, y \geq 1 \end{cases}$$

توزیع‌های حاشیه‌ای - مثال

دو قرص به تصادف از بسته‌ای محتوای سه قرص اسپرین و ۲ قرص خواب‌آور و ۴ قرص سرماخوردگی انتخاب می‌کنیم.

X تعداد قرص اسپرین، Y تعداد قرص خواب‌آور

توزیع احتمال X و توزیع احتمال Y

X	Y
0	0
0	۱
۱	0
۱	۱
0	۲
۲	0

	$X = 0$	$X = 1$	$X = 2$	
$Y = 0$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
$Y = 1$	$\frac{2}{9}$	$\frac{1}{6}$		$\frac{7}{18}$
$Y = 2$	$\frac{1}{36}$			$\frac{1}{36}$
	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	

توزیع‌های حاشیه‌ای - گسسته

X و Y دو متغیر تصادفی گسسته

$f(x,y)$ توزیع احتمال توام آنها

▪ توزیع حاشیه‌ای $X=x$

$$f_X(x) = \sum_y f_{XY}(x, y)$$

▪ توزیع حاشیه‌ای $Y=y$

$$f_Y(y) = \sum_x f_{XY}(x, y)$$

توزیع‌های حاشیه‌ای-پیوسته

X و Y دو متغیر تصادفی پیوسته

$f(x,y)$ چگالی احتمال توام آنها

▪ چگالی حاشیه‌ای $X=x$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy$$

▪ چگالی حاشیه‌ای $Y=y$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx$$

چگالی حاشیه‌ای - مثال

چگالی توام

$$f_{XY}(x, y) = \begin{cases} \frac{2}{3}(x + 2y), & 0 < x < 1, 0 < y < 1 \\ 0, & \text{در غیر این صورت} \end{cases}$$

چگالی‌های حاشیه‌ای X و Y

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^1 \frac{2}{3}(x + 2y) dy = \frac{2}{3}(x + 1)$$


$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_0^1 \frac{2}{3}(x + 2y) dx = \frac{1}{3}(1 + 4y)$$

توزیع‌های توام و حاشیه‌ای - تابع چگالی

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY}(x, y) dy dx = 1$$

چگالی احتمال حاشیه‌ای X

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy$$


$$\sum_{x,y} f_{XY}(x, y) = \sum_x \sum_y f_{XY}(x, y)$$

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{s \leq x} \sum_{t \leq y} \underbrace{f_{XY}(s, t)}_{\text{جرم}}$$

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^x \underbrace{f_{XY}(s, t)}_{\text{چگالی}} ds dt$$

چند متغیری گسسته

$$f(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

مثال

$$f(x,y,z) = \frac{(x+y)z}{6^3}, x = 1,2; y = 1,2,3; z = 1,2$$

$$?P(X = 2, Y + Z \leq 3)$$

$$\begin{aligned} P(X = 2, Y + Z \leq 3) &= f(2, 1, 1) + f(2, 1, 2) + f(2, 2, 1) \\ &= \frac{3}{6^3} + \frac{6}{6^3} + \frac{4}{6^3} = \frac{13}{6^3} \end{aligned}$$

چند متغیری پیوسته

احتمال با استفاده از چگالی توام

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n$$

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n}{\partial x_1 \partial x_2 \dots \partial x_n} F(x_1, x_2, \dots, x_n)$$

توزیع‌های شرطی

با دانستن احتمال X در مقدار x ، توزیع احتمال Y چقدر است؟
▪ گسسته - تابع جرم

$$P_{Y|X}(y|x) = \frac{P_{XY}(x, y)}{P_X(x)}$$

▪ پیوسته - تابع چگالی

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

توزیع‌های شرطی - مثال

	$X = 0$	$X = 1$	$X = 2$	
$Y = 0$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
$Y = 1$	$\frac{2}{9}$	$\frac{1}{6}$		$\frac{7}{18}$
$Y = 2$	$\frac{1}{36}$			$\frac{1}{36}$
	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	

دو قرص به تصادف از بسته‌ای محتوای سه قرص اسپرین و ۲ قرص خواب‌آور و ۴ قرص سرماخوردگی انتخاب می‌کنیم.

X تعداد قرص اسپرین، Y تعداد قرص خواب‌آور

$$P(X = 0 | Y = 1) = \frac{\frac{2}{9}}{\frac{7}{18}} = \frac{4}{7}$$

$$P(X = 1 | Y = 1) = \frac{\frac{1}{6}}{\frac{7}{18}} = \frac{3}{7}$$

$$P(X = 2 | Y = 1) = \frac{0}{\frac{7}{18}} = 0$$

$$f_{XY}(x, y) = \begin{cases} \frac{2}{3}(x + 2y), & 0 < x < 1, 0 < y < 1 \\ 0, & \text{در غیر این صورت} \end{cases}$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^1 \frac{2}{3}(x + 2y) dy = \frac{2}{3}(x + 1)$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_0^1 \frac{2}{3}(x + 2y) dx = \frac{1}{3}(1 + 4y)$$

مثال

چگالی شرطی X با داشتن $Y = y$: $P\left(x \leq \frac{1}{2} \mid y = \frac{1}{2}\right)$

$$f(x|y) = \frac{f(x, y)}{f(y)} = \frac{\frac{2}{3}(x + 2y)}{\frac{1}{3}(1 + 4y)} = \begin{cases} \frac{(2x + 4y)}{(1 + 4y)}, & 0 < x < 1 \\ 0, & \text{در غیر این صورت} \end{cases}$$

$$f\left(x \mid y = \frac{1}{2}\right) = \frac{(2x + 4y)}{(1 + 4y)} = \frac{(2x + 4 \times \frac{1}{2})}{(1 + 4 \times \frac{1}{2})} = \frac{2x + 2}{3}$$

$$P\left(x \leq \frac{1}{2} \mid y = \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \frac{2x + 2}{3} dx = \frac{5}{12}$$

توزیع شرطی چند متغیری

دارای انواع متفاوت

اهمیت استقلال

توزیع‌های توام و حاشیه‌ای - استقلال

م‌ت‌های X و Y مستقلند اگر $F_{XY}(x,y) = F_X(x) F_Y(y)$

▪ گسسته: برای تمامی

$$p_{XY}(x, y) = p_X(x)p_Y(y)$$

$$p_{Y|X}(y | x) = p_Y(y)$$

▪ پیوسته: برای تمام اعداد حقیقی x, y

$$f_{XY}(x, y) = f_X(x)f_Y(y)$$

$$f_{Y|X}(y | x) = f_Y(y)$$

توزیع‌های توام و حاشیه‌ای - استقلال

لم- اگر X و Y مستقل باشند، آن‌گاه هر زیرمجموعه A و B از اعداد حقیقی خواهیم داشت:

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

نشان می‌دهد که هر تابعی از X ، از هر تابعی از Y مستقل خواهد بود.

توزیع‌های بیز

با دانستن احتمال X در مقدار x ، توزیع احتمال Y چقدر است؟
▪ گسسته - تابع جرم

$$P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)} = \frac{P_{X|Y}(x|y)P_Y(y)}{\sum_{y'} P_{X|Y}(x|y')P_Y(y')}$$

▪ پیوسته - تابع چگالی

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{f_{X|Y}(x|y)f_Y(y)}{\int_{-\infty}^{+\infty} f_{X|Y}(x|y')f_Y(y')dy'}$$

قاعده زنجیری

متغیرهای X_1, \dots, X_k

$$P_{X_1, \dots, X_k}(x_1 \cap x_2 \cap \dots \cap x_k)$$

$$= P_{X_1}(x_1) P_{X_2|X_1}(x_2|x_1) \dots P_{X_k|X_1, \dots, X_{k-1}}(x_k|x_1 \cap x_2 \cap \dots \cap x_{k-1})$$

امید ریاضی

در ارتباط با بازی‌های شانسی
▪ صورت معمول آن معادل مبلغی حاصل از بازی بازیکن

گسسته

$$\mathbb{E}[X] = \sum_x x P_X(x)$$

پیوسته

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

مثال

۱۲ تلویزیون شامل ۲ تلویزیون معیوب
انتظار دستگاه معیوب با انتخاب ۳ دستگاه

$$f(x) = \frac{\binom{2}{x} \binom{10}{3-x}}{\binom{12}{3}}, \quad x = 0, 1, 2$$

x	0	1	2
$f(x)$	$\frac{6}{11}$	$\frac{9}{22}$	$\frac{1}{22}$

$$E(X) = 0 \cdot \frac{6}{11} + 1 \cdot \frac{9}{22} + 2 \cdot \frac{1}{22}$$

مثال

$$f(x) = \begin{cases} \frac{4}{\pi(1+x^2)}, & 0 < x < 1 \\ 0 & \text{سایر جاها} \end{cases}$$

$$E(X) = \int_0^1 x \frac{4}{\pi(1+x^2)} = \frac{4}{\pi} \int_0^1 \frac{x}{1+x^2}$$

امید ریاضی

$$g: \mathbb{R} \rightarrow \mathbb{R}$$

گسسته

$$\mathbb{E}[g(X)] = \sum_x g(x)P_X(x)$$

پیوسته

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx$$

حالت خاص: $g(X) = X$

مثال

X عدد ریختن تاس

$$g(X) = 2X^2 + 1$$

$$\begin{aligned}\mathbb{E}[g(X)] &= \sum_{x=1}^6 (2x^2 + 1) \cdot \frac{1}{6} \\ &= (2 \times 1^2 + 1) \cdot \frac{1}{6} + \dots + (2 \times 6^2 + 1) \cdot \frac{1}{6} = \frac{94}{3}\end{aligned}$$

امید ریاضی - ویژگی‌ها

$$a \in \mathbb{R}$$

$$\mathbb{E}[a] = a$$

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

$$\mathbb{E}[af(X)] = a\mathbb{E}[f(X)]$$

خطی بودن

$$\mathbb{E}[f(X) + g(X)] = \mathbb{E}[f(X)] + \mathbb{E}[g(X)]$$

متغیر تصادفی گسسته X

$$\mathbb{E}[1_{\{X = k\}}] = P(X = k)$$

وردائی یا واریانس

سنجش پراکندگی

چگونگی تمرکز مت X حول میانگین

$$\begin{aligned} \text{Var}[X] &= E[(X - E[X])^2] \\ &= E[X^2] - E[X]^2 \end{aligned}$$

وردائی یا واریانس - ویژگی ها

$$\text{Var}[a] = 0$$

$$\text{Var}[af(X)] = a^2 \text{Var}[f(X)]$$

وردائی یا واریانس - مثال

میانگین و وردائی متغیر تصادفی یکنواخت X با چتا

$$f_X(x) = \begin{cases} 1, \forall x \in [0,1] \\ 0, \forall x \notin [0,1] \end{cases}$$

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_0^1 x dx = \frac{1}{2}$$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \int_0^1 x^2 dx = \frac{1}{3}$$

$$Var[X] = E[X^2] - E[X]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

مثال

فرض کنید $g(x) = \mathbf{1}\{x \in A\}$ به ازای $A \subseteq \Omega$ و $E[g(X)]$ ؟

حالت گسسته

$$E[g(X)] = \sum_{x \in X} \mathbf{1}\{x \in A\} P_X(x) = \sum_{x \in A} P_X(x) = P_X(x \in A)$$

حالت پیوسته

$$E[g(X)] = \int_{-\infty}^{+\infty} \mathbf{1}\{x \in A\} f_X(x) dx = \int_{x \in A} f_X(x) dx = P_X(x \in A)$$

چبیشف کبیر!

اگر μ میانگین و σ انحراف معیار متغیر تصادفی X باشند، آن گاه برای هر مقدار مثبت k داریم

$$P(|x - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

اثبات

$$\sigma^2 = E[(X - \mu)^2]$$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$

$$\sigma^2 = \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (x - \mu)^2 f(x) dx$$

انتگرال $(x - \mu)^2 f(x)$ نامنفی \Leftarrow پس

$$\sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (x - \mu)^2 f(x) dx$$

چون $x \leq \mu - k\sigma$ یا $x \geq \mu + k\sigma$ پس $(x - \mu)^2 \geq k^2 \sigma^2$. در نتیجه

$$\sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} k^2 \sigma^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} k^2 \sigma^2 f(x) dx$$

اثبات

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} k^2 \sigma^2 f(x) dx + \int_{\mu+k\sigma}^{+\infty} k^2 \sigma^2 f(x) dx$$

با شرط $\sigma^2 \neq 0$

$$\frac{1}{k^2} \geq \int_{-\infty}^{\mu-k\sigma} f(x) dx + \int_{\mu+k\sigma}^{+\infty} f(x) dx$$

پس

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

و یا

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

چبیشف

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

من باب مثال

- احتمال قرار گرفتن X در فاصله 2σ میانگین حداقل $1 - \frac{1}{2^2} = \frac{3}{4}$
- احتمال قرار گرفتن X در فاصله 3σ میانگین حداقل $1 - \frac{1}{3^2} = \frac{8}{9}$
- احتمال قرار گرفتن X در فاصله 5σ میانگین حداقل $1 - \frac{1}{5^2} = \frac{24}{25}$

σ کنترل کننده پراکندگی توزیع متغیر تصادفی

مثال

$$f(x) = \begin{cases} 63 \cdot x^4 (1-x)^4, & 0 < x < 1 \\ 0 & \text{سایر جاها} \end{cases}$$

احتمال قرار گرفتن X در فاصله 2σ از میانگین و مقایسه با کران پائینی حاصل از قضیه چبیشف

$$\sigma = \sqrt{\frac{1}{44}} = 0.15 \leftarrow \sigma^2 = \frac{1}{44} \text{ و میانگین } \mu = \frac{1}{2}$$

احتمال بودن در فاصله مدنظر در بازه 0.2 و 0.8

$$P(0.2 < X < 0.8) = \int_{0.2}^{0.8} 63 \cdot x^4 (1-x)^4 dx = 0.96$$

نتیجه حاصل از قضیه چبیشف 0.75

امید ریاضی

گسسته

$$\mathbb{E}[XY] = \sum_x XY P_{XY}(x, y)$$

پیوسته

$$\mathbb{E}[XY] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} XY f_{XY}(x, y) dx dy$$

$$\text{؟ } E[XY] = E[X]E[Y]$$

امید ریاضی

$$g: \mathbb{R}^2 \rightarrow \mathbb{R}$$

گسته

$$\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y) P_{XY}(x, y)$$

پیوسته

$$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{XY}(x, y) dx dy$$

توزیع‌های توام و حاشیه‌ای - امید ریاضی و کوواریانس

استفاده از امید ریاضی جهت مطالعه روابط بین دو مت

$$\text{Cov}[X, Y] = E[(X - E(X))(Y - E(Y))] = E[XY] - E[X]E[Y]$$

$$? \text{Cov}[X, Y] = 0$$

توزیع‌های توأم و حاشیه‌ای - امید ریاضی و کوواریانس

ویژگی‌ها

$$E[f(X, Y) + g(X, Y)] = E[f(X, Y)] + E[g(X, Y)]$$

$$E[aX + bY] = aE[X] + bE[Y]$$

$$Var[X + Y] = Var[X] + Var[Y] + 2Cov[X, Y]$$

اگر X و Y مستقل باشند، آن‌گاه

$$E[XY] = E[X]E[Y]$$

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)]$$

$$Cov[X, Y] = 0$$

توزیع‌های حاشیه‌ای - مثال

		$X = 0$	$X = 1$	$X = 2$	
	$Y = 0$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
	$Y = 1$	$\frac{2}{9}$	$\frac{1}{6}$		$\frac{7}{18}$
	$Y = 2$	$\frac{1}{36}$			$\frac{1}{36}$
		$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	
X	Y				
0	0				
0	1				
1	0				
1	1				
0	2				
2	0				

دو قرص به تصادف از بسته‌ای محتوای سه قرص اسپرین و ۲ قرص خواب‌آور و ۴ قرص سرماخوردگی انتخاب می‌کنیم.

X تعداد قرص اسپرین، Y تعداد قرص خواب‌آور

توزیع احتمال X و توزیع احتمال Y

$$\mu_{XY} = E(XY) = 0 \times 0 \times \frac{1}{6} + 0 \times 1 \times \frac{2}{9} + 0 \times 2 \times \frac{1}{36} + 1 \times 0 \times \frac{1}{3} + 1 \times 1 \times \frac{1}{6} + 2 \times 0 \times \frac{1}{12} = \frac{1}{6}$$

$$\mu_X = E(X) = 0 \times \frac{5}{12} + 1 \times \frac{1}{2} + 2 \times \frac{1}{12} = \frac{2}{3}$$

$$\mu_Y = E(Y) = 0 \times \frac{7}{12} + 1 \times \frac{7}{18} + 2 \times \frac{1}{36} = \frac{4}{9}$$

$$\sigma_{XY} = \frac{1}{6} - \frac{2}{3} \times \frac{4}{9} = -\frac{7}{54}?$$

توزیع‌های توأم و حاشیه‌ای - امید ریاضی و کوواریانس

ویژگی‌ها

$$\text{Cov}[aX + b, cY + d] = ac\text{Cov}(X, Y)$$

$$\text{Cov}[X, Y] = \text{Cov}[Y, X]$$

$$\text{Cov}[X, Y + Z] = \text{Cov}[Y, X] + \text{Cov}(X, Z)$$

$$\text{Cov}[X, a] = 0$$

توزیع‌های توأم و حاشیه‌ای - امید ریاضی و کوواریانس

ویژگی‌ها

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$$

استقلال دو متغیر

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$

$$\text{Var}[aX + bY] = a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab\text{Cov}(X, Y)$$

مثال

سه متغیر تصادفی X و Y و Z

$$\mu_X = 2, \mu_Y = -3, \mu_Z = 4$$

$$\sigma_X^2 = 1, \sigma_Y^2 = 5, \sigma_Z^2 = 2$$

$$\text{Cov}(X, Y) = -2, \text{Cov}(X, Z) = -1, \text{Cov}(Y, Z) = 1$$

میانگین و وردائی $W = 3X - Y + 2Z$

$$E(W) = E(3X - Y + 2Z) = 3E(X) - E(Y) + 2E(Z) = 3 \times 2 - (-3) + 2 \times 4 = 17$$

$$\text{Var}(W) = 9\text{Var}(X) + \text{Var}(Y) + 4\text{Var}(Z) - 6\text{Cov}(X, Y) + 12\text{Cov}(X, Z) - 4\text{Cov}(Y, Z)$$

$$= 9 \times 1 + 5 + 4 \times 2 - 6 \times (-2) + 12(-1) - 4 \times 1 = 18$$

چند توزیع پر استفاده

بررسی میانگین و وردائی

توزیع های احتمال گسسته

چگالی های احتمال پیوسته

توزیع یکنواخت

برآمد تمام خروجی‌ها دارای احتمال برابر

$$f(x) = \frac{1}{k}, x = x_1, \dots, x_k$$

توزیع برنولی

آزمایش دارای دو برآمد پیروزی و شکست، و احتمال متناظر θ و $1 - \theta$

$X \sim (\theta)$ برنولی

$$0 \leq \theta \leq 1$$

شیر=1

خط=0

$$f(x; \theta) = \theta^x (1 - \theta)^{1-x}, x = 0, 1$$

توزیع دوجمله‌ای

امتحان‌های تکراری

احتمال تمام پیروزی‌ها برابر با θ

هر پرتاب (آزمایش) مستقل از دیگر آزمایش‌ها

$X \sim \text{دوجمله‌ای}(x; n, \theta)$

$$0 \leq \theta \leq 1$$

$$b(x; n, \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \quad x = 0, 1, \dots, n$$

مثال

بهبود ۷ نفر از ۱۰ نفر با احتمال بهبود ۰.۸

$$b(7; 10, 0.8) = \binom{10}{7} 0.8^7 (1 - 0.8)^3 \approx 0.2$$

ویژگی‌ها

احتمال تجمیعی

$$B(x; n, \theta) = \sum_{k=0}^x b(x; n, \theta)$$

$$b(x; n, \theta) = b(n - x; n, 1 - \theta)$$

$$\mu = n\theta$$

$$\sigma^2 = n\theta(1 - \theta)$$

توزیع فوق هندسی

نمونه گیری بدون جایگذاری

- وابستگی آزمایش ها به یکدیگر
- فضای نمونه شامل N عضو
- M پیروزی و $N-M$ شکست
- انتخاب n عضو بدون جایگذاری
- احتمال x پیروزی از n انتخاب

$X \sim$ فوق هندسی $(x; n, N, M)$

$$h(x; n, N, M) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}},$$

$$x = 0, 1, 2, 3, \dots; x \leq M; n - x \leq N - M$$

ویژگی‌ها

$$\mu = \frac{nM}{N}$$

$$\sigma^2 = \frac{nN(N-M)(N-n)}{N^2(N-1)}$$

با شرط N بزرگ و $n \ll N$

▪ عدم تفاوت بین نمونه‌گیری با جاگذاری و بدون جای‌گذاری

▪ امکان استفاده از تقریب احتمال فوق هندسی $h(x; n, N, M)$ با معادله توزیع دو جمله‌ای $b\left(x; n, \frac{M}{N}\right)$

توزیع پواسن

سختی محاسبه احتمال دوجمله‌ای $b(x; n, \theta)$ با بزرگ بودن n

$X \sim \text{پواسن}(x; \lambda)$

میانگین $\lambda > 0$

توزیع احتمال حول عدد صحیح نامنفی که برای مدل‌سازی فراوانی یا بسامد پیشامدهای نادر استفاده دارد.

$$p(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots$$

توزیع چند جمله‌ای

$$f(x_1, x_2, \dots, x_k; n, \theta_1, \theta_2, \dots, \theta_k) = \binom{n}{x_1, x_2, \dots, x_k} \theta_1^{x_1} \theta_2^{x_2} \dots \theta_k^{x_k}$$
$$x_1 + x_2 + \dots + x_k = n$$
$$\theta_1 + \theta_2 + \dots + \theta_k = 1$$

توزیع فوق هندسی چندمتغیره

$$f(x_1, x_2, \dots, x_k; n, M_1, M_2, \dots, M_k) = \frac{\binom{M_1}{x_1} \binom{M_2}{x_2} \cdots \binom{M_k}{x_k}}{\binom{N}{n}}$$

$$x_i = 0, 1, 2, \dots, n$$

$$x_i \leq M_i$$

$$x_1 + x_2 + \cdots + x_k = n$$

$$M_1 + M_2 + \cdots + M_k = N$$

توزیع‌های پیوسته

توزیع یکنواخت

$X \sim$ یکنواخت (a, b)

$a < b$

چگالی احتمال یکسان برای هر نقطه بین a و b

$$u(x; a, b) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0 & \text{سایر جاها} \end{cases}$$

چگالی احتمال گاما

$$g(x; a, b) = \begin{cases} \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-\frac{x}{b}}, x > 0 \\ 0 \text{ ساير جاها} \end{cases}$$

چگالی احتمال نمائی

$$\theta > 0$$

چگالی احتمال کاهشی روی مقادیر حقیقی نامنفی

$$g(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, x > 0 \\ 0 \text{ سایر جاها} \end{cases}$$

چگالی احتمال بتا

$$g(x; a, b) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}, & 0 < x < 1 \\ 0 & \text{سایر جاها} \end{cases}$$

$$a, b > 0$$

چگالی احتمال نرمال (گوسی یا زنگوله‌ای)

$$n(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < +\infty, \sigma > 0$$

توزیع نرمال استاندارد

$$\mu = 0, \sigma = 1 \quad \blacksquare$$

ارزش عقیده دوم

به دلیل مهلك بودن بیماری، به دنبال بررسی دومی درباره راستی آزمایی وجود بیماری

$$P(\beta|+test) = \frac{P(+test|\beta)P(\beta)}{P(+test|\beta)P(\beta) + P(+test|\neg\beta)P(\neg\beta)}$$

پزشک دیگر

▪ نوع دیگر و مستقلی از آزمایش یا تفسیر و خوانش دیگری از آزمایش یکسان با قبلی

▪ اغلب خوانش چندباره یک آزمایش غالب تر بر آزمایش های دیگر. چرا؟

▪ نتیجه یا مثبت (تائید) یا منفی

$$P(+test|\neg\beta) = 0.03$$

$$P(-test|\beta) = 0.01$$

$$P(+test|\beta) = ? = 1 - P(-test|\beta) = 1 - 0.01 = 0.99$$

$$P(\beta) = 0.032$$

فرض نتیجه مثبت!

فرض مثبت غلط و منفی غلط یکسان با قبل

▪ ابتدا نتیجه 0.03%

▪ سپس با اطلاع از بیماری خانوادگی 3.2%

▪ آخرین احتمال به عنوان احتمال پیشین

▪ نتیجه 52.2%

$$P(\beta|+test) = \frac{0.99 \times 0.032}{0.99 \times 0.032 + 0.03 \times (1 - 0.032)} = 0.522$$

منابع تهیه این درس

[پینسکی]

ج. فروند و همکاران، «آمار ریاضی و کاربردهای آن»، م. ق. وحیدی اصل، ع. عمیدی، مرکز نشر دانشگاهی، چاپ دوم، ۱۳۹۲ - فصل‌های ۲ و ۳ و بخشی از فصل ۴

[پرس] S. J. Press, J. M. Tanur, "The Subjectivity of Scientists and the Bayesian Approach", Wiley, 2001

CS 228 - Probabilistic Graphical Models, 2018-19, Stanford University,
<https://cs228.stanford.edu>