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

رابطهٔ دقیق بین ورودی و خروجی با مقدار دیگری تدوین می‌شود که نمایشگر این است که ورودی چقدر بر خروجی تأثیر می‌گذارد

- پارامتر

تدوین مدل (معادله‌بندی مدل): نمایشگر اینکه ورودی بر خروجی تأثیرگذار است.

- پارامترها مقدار دقیق تاثیر را معین می‌کنند

مدل دارای پارامتر است

مثال

- شیر آمدن سکه
  - سکهٔ 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 پارامترهائی است.

مقدار پارامترها مشخص‌کنندۀ 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 دربارۀ 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 واکسن انفلوانزا هنگام واکسن 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 مطالعهٔ دقیقِ عدم‌قطعیت‌ها

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 عدم‌قطعیت‌ها

نسبتی از زمان اتفاق افتادن پیشامد

درجه اعتقاد درباره پیشامد

عدم‌قطعیت 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
- سکهٔ سالم $\theta = 0.5$
- درجهٔ اعتقاد $\mathrm{P}(\theta)$
    - $\mathrm{P}(\theta = 0.5) = 0.99$
- احتمال:
    - شیر و خط
    - فضای نمونه دارند
- درجهٔ اعتقاد
    - مقدار پیوسته $\theta =$ ۰و۰.۰۱و۰.۰۲و ... و۱.۰۰

چرا سکه

- چقدر سکه پرتاب می‌کنیم؟
- مشابه زندگی- احتمال زنده ماندن یکسال پس از جراحی
- احتمال اثرجانبی 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 اندازه‌گیرندهٔ 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 ترکیباتی

اعمال چند مرحله‌ای

- دومرحله‌ای
- بیش از دو مرحله

جایگشت (آرایش)

- n شی متمایز
- چینش r شی از n شی متمایز
- n شی متمایز روی دایره
- n شی به طوری که $n_1 + n_2 + \cdots + n_k = n$

ترکیب (زیرمجموعه)

- انتخاب r شی از n شی متمایز
- تعداد افرازهای n شی متمایز به r زیرمجموعه با تعداد معین

ضرائب دوجمله‌ای

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

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

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

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

معادلهٔ استرلینگ

$$n! \cong \sqrt{۲\pi n}\left(\frac{n}{e}\right)^n$$

ضرائب چندجمله‌ای

$$(x_۱ + x_۲ + \cdots + x_k)^n$$
$$x_۱^{r_۱} x_۲^{r_۲} x_k^{r_k}, \binom{n}{r_۱, r_۲, \cdots, r_k}$$

ضرائیب دوجمله‌ای تعیمی یافته

$$\binom{n}{r} = \frac{n(n-1)\ldots(n-r+۱)}{r!}$$

مثال $\binom{-۱}{r}$؟

عدد اویلر

$$\underset{n\to\infty}{\text{حــــد}}\left(1+\frac{1}{n}\right)^{n}=?$$

$$\underset{n\to\infty}{\text{حــــد}}\left(1+\frac{a}{n}\right)^{n}=?$$

$$\underset{n\to\infty}{\text{حــــد}}\left(1-\frac{a}{n}\right)^{n}=?$$

تفسیرهای احتمال

کلاسیک (احتمال برابر)

بسامدی

بیزی

اصل آغازه

فضای نمونه Ω

آزمایش
- هر فرایند مشاهده یا اندازه‌گیری

برآمد
- نتایج حاصل از آزمایش

فضای نمونه
- مجموعهٔ 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 وسیلهٔ الکترونیکی

- فضای نمونه

وسیلهٔ الکترونیکی قبل از ششمین سال از کار بیفتد

- پیشامد

$$\Omega = \{\mathrm{t}: \mathrm{t} \geq 0\}$$
$$F = \{t: 0 \leq t \leq 6\}$$

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 پیشامدA 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 تک‌تک 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

$$p(\phi) = 0$$

$$A \subseteq B \Longrightarrow P(A) \leq P(B)$$
$$0 \leq P(A) \leq 1$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cup B \cup C)$$
$$= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

ویژگی‌ها

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

$$P(A \cap B) \leq \min(P(A), P(B))$$
$$p(\phi) = 0$$

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 احتمال پیشامد A است بعد از مشاهدهء پیشامد B

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

مثال

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

خ: احتمال مرکز مناسب چقدر است؟

$$P(\text{خ}) = \frac{۲۶}{۵۰} = ۰.۵۲$$

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

$$p\left(\text{خ}\middle|\text{ک}\right) = \frac{p\left(\text{ک} \cap \text{خ}\right)}{p\left(\text{ک}\right)} = \frac{\frac{۱۶}{۵۰}}{\frac{۲۰}{۵۰}} = \frac{۱۶}{۲۰} = ۰.۸$$

قاعدۀ زنجیری

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

برای $k = ۲$

$$P(S_۱ \cap S_۲) = P(S_۱)P(S_۲|S_۱)$$

در واقع قاعدهء زنجیری مشتق از اعمال چندبارهء احتمال شرطی است.

برای $k = ۳$

$$P(S_۱ \cap S_۲ \cap S_۳) = P(S_۱)P(S_۲|S_۱)P(S_۳|S_۱ \cap S_۲)$$

برای k

$$P(S_۱ \cap S_۲ \cap \cdots \cap S_k) = P(S_۱)P(S_۲|S_۱)P(S_۳|S_۱ \cap S_۲) \ldots P(S_k|S_۱ \cap S_۲ \cap \cdots \cap S_{k-۱})$$

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

$$p\left(\text{خ} \cap \text{ک}\right) = p\left(\text{خ}\right) p\left(\text{ک} \middle| \text{خ}\right) = \frac{۱۵}{۲۴۰}\frac{۱۴}{۲۳۹}$$

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؟

$$p\left(\text{خ} \cap \text{ک} \cap \text{ل}\right) = p\left(\text{خ}\right) p\left(\text{ک} \middle| \text{خ}\right) p\left(\text{ل} \middle| \text{خ} \cap \text{ک}\right) = \frac{\text{۵}}{\text{۲۰}} \frac{\text{۴}}{\text{۱۹}} \frac{\text{۳}}{\text{۱۸}}$$

استقلال

دو پیشامد مستقل خوانده می‌شوند اگر

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

یا

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

شهود دو پیشامد مستقل

- عدم‌تأثیر مشاهدهٔ B بر مشاهدهٔ A

مثال

۲۰ فیوز در بسته داریم که ۵ مورد انها 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؟

$$p\left(\text{خ}\right)=\frac{۵}{۲۰}$$

$$p\left(\text{ک}\middle|\text{خ}\right)=\frac{۵}{۲۰}$$

$$p\left(\text{ل}\middle|\text{خ}\cap\text{ک}\right)=\frac{۵}{۲۰}$$

$$p\left(\text{خ}\cap\text{ک}\cap\text{ل}\right)=p\left(\text{خ}\right)p\left(\text{ک}\middle|\text{خ}\right)p\left(\text{ل}\middle|\text{خ}\cap\text{ک}\right)=(\frac{۵}{۲۰})^{۳}$$

مثال

سکه‌ای را سه بار پرتاب کنیم. خ: ۱ ش: ۰، $\Omega = \{000,001,010, \dots ,111\}$

احتمال پیشامد شیر در دو پرتاب اول α، $p(\alpha) = \frac{۱}{۴}$

احتمال پیشامد خط در پرتاب سوم β، $p(\beta) = \frac{۱}{۲}$

احتمال پیشامد دقیقا دو خط در سه پرتاب γ، $p(\gamma) = \frac{۳}{۸}$

پیشامدهای α و β مستقل‌اند؟ $p(\alpha\beta) = \frac{۱}{۸} = p(\alpha)p(\beta)$

پیشامدهای γ و β مستقل‌اند؟ $p(\gamma \cap \beta) = \frac{۱}{۴}, p(\gamma)p(\beta) = \frac{۳}{۱۶}$

قضیهٔ بیز

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.

$$p(\beta) = ۰.۶$$

$$p(\alpha|\beta) = ۰.۳۵$$

$$p(\alpha|\beta') = ۰.۸۵$$

$$p(\alpha) = p\big((\alpha \cap \beta) \cup (\alpha \cap \beta')\big)$$

$$= p(\alpha \cap \beta) + p(\alpha \cap \beta')$$

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

$$= ۰.۶ \times ۰.۳۵ + ۰.۴ \times ۰.۸۵ = ۰.۵۵$$

قاعدهٔ احتمال کل (قاعدهٔ حذف)

پیشامدهای β_1، β_2، …، β_k افرازی از فضای نمونهٔ 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؟

اتومبیل دودزا $p(\beta) = ۰.۲۵$

اتومبیل سالم $p(\beta') = ۰.۷۵$

رد شدن α

$$p(\alpha|\beta) = ۰.۹۹$$

$$p(\alpha|\beta') = ۰.۱۷$$

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

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

قاعدهٔ بیز

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

تعمیم

پیشامدهای $\beta_۱$، $\beta_۲$، ...، β_k افرازی از فضای نمونهٔ 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 صرفا توجه به مجموعه رقم‌ها
  - (1,1)
  - (2,1)
  - ...

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

فضای هدف

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

- متغیر تصادفی
- غلط مصطلح و سوتفاهم
  - 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 ده‌تائی 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.
- مثلا مجموع دو عددی که در ریختن جفت تاس، یا تعداد شیرهائی که در ۱۰ پرتاب تاس به دست می‌آید

مت:

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

نمایش با X(ω) یا X

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

مثال

سکه‌ای چهار با پرتاب می‌شود. مت تعداد کل شیرها

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

$$P(X = ۳) = \frac{۴}{۱۶}$$

فضای نمونه	احتمال	X=x
0000	$\frac{۱}{۱۶}$	۴
0001	$\frac{۱}{۱۶}$	۳
0010	$\frac{۱}{۱۶}$	۳
0011	$\frac{۱}{۱۶}$	۲
0100	$\frac{۱}{۱۶}$	۳
0101	$\frac{۱}{۱۶}$	۲
0110	$\frac{۱}{۱۶}$	۲
0111	$\frac{۱}{۱۶}$	۱
1000	$\frac{۱}{۱۶}$	۳
1001	$\frac{۱}{۱۶}$	۲
1010	$\frac{۱}{۱۶}$	۲
1011	$\frac{۱}{۱۶}$	۱
1100	$\frac{۱}{۱۶}$	۲
1101	$\frac{۱}{۱۶}$	۱
1111	$\frac{۱}{۱۶}$	۰

متغیرهای تصادفی - ادامه

اگر Ω فضای نمونه با اندازه احتمال باشد 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 وسیلهء الکترونیکی

- مقداری نامتناهی را می‌پذیرد
- b و a به طوری که $a < b$

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

مثال- توزیع احتمال

فضای نمونهٔ پرتاب جفت تاس هر عضو فضا $\frac{۱}{۳۶}$

متغیر تصادفی حاصل‌جمع دو عدد تاس‌ها

تدوین معادله

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

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	$\frac{۱}{۳۶}$

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 روزانهٔ 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

توزیع‌های احتمال

بر اساس اصول کولمولگروف

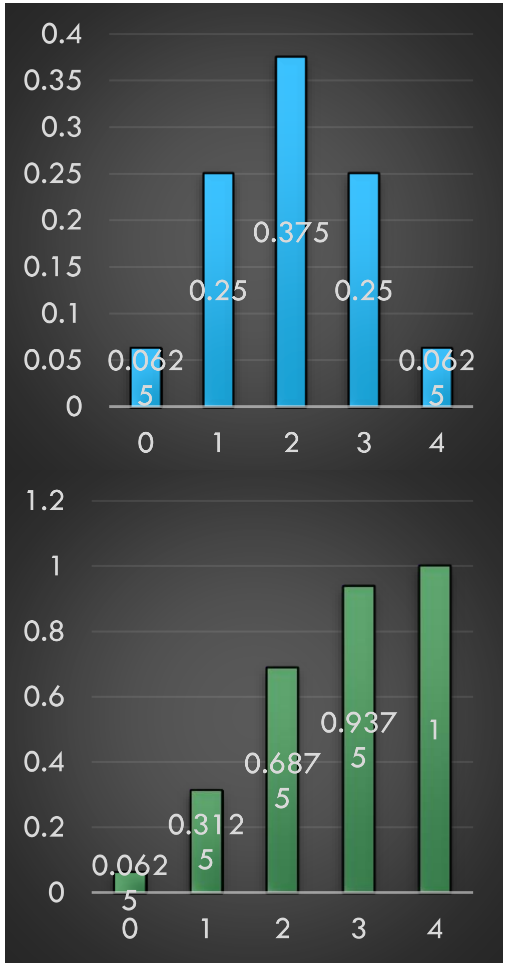
قضیه: توابعی را می‌توان به مثابهٔ توزیع احتمالیِ متغیر تصادفی گسسته به کار برد که مقادیر $f(x)$ اگر و فقط اگر مشمول شرایط زیر باشند

$$f(x) \geq 0$$

$$\sum_{x} f(x) = 1$$

مثال- نمودار میله‌ای

$$P(X=0)=\frac{۱}{۱۶}$$
$$P(X=۱)=\frac{۴}{۱۶}$$
$$P(X=۲)=\frac{۶}{۱۶}$$
$$P(X=۳)=\frac{۴}{۱۶}$$
$$P(X=۴)=\frac{۱}{۱۶}$$



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.

$$f(x) = \frac{x + ۲}{۲۵}, x = ۱, ۲, ۳, ۴, ۵$$

$$f(x) \geq 0$$

$$\sum_{x} f(x) = 1$$

توابع توزیع تجمیعی-

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

مثال

توزیع احتمال این تابع تجمیعی چقدر است؟

x	$f(X=x)$
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	$\frac{۱}{۳۶}$

$$F(x) = \begin{cases} 0, & x < ۲ \\ \frac{۱}{۳۶}, & ۲ \leq x < ۳ \\ \frac{۳}{۳۶}, & ۳ \leq x < ۴ \\ \frac{۶}{۳۶}, & ۴ \leq x < ۵ \\ & \vdots \\ \frac{۳۵}{۳۶}, & ۱۱ \leq x < ۱۲ \\ ۱, & x \geq ۱۲ \end{cases}$$

$$P(۲) = \frac{۱}{۳۶}$$

$$P(۳) = \frac{۳}{۳۶} - \frac{۱}{۳۶} = \frac{۲}{۳۶}$$

$$P(۴) = \frac{۶}{۳۶} - \frac{۳}{۳۶} = \frac{۳}{۳۶}$$

$$\vdots$$

$$P(۱۲) = ۱ - \frac{۳۵}{۳۶} = \frac{۱}{۳۶}$$

توابع توزیع تجمیعی - ادامه

ویژگی‌ها

$$0 \leq F(X) \leq ۱$$

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

$$\lim_{x \to +\infty} F(X) = ۱$$

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

توابع توزیع تجمیعی - ادامه

قضیه: اگر متغیر تصادفی متشکل از مقادیر $x_n > \cdots > x_۲ > x_۱$ باشد، آن‌گاه $F(x_۱) = \mathrm{P}(x_۱)$، و

$$P(x_i) = F(x_i) - F(x_{i-1}) \ , i = ۲, ۳, \ldots, n$$

مثال

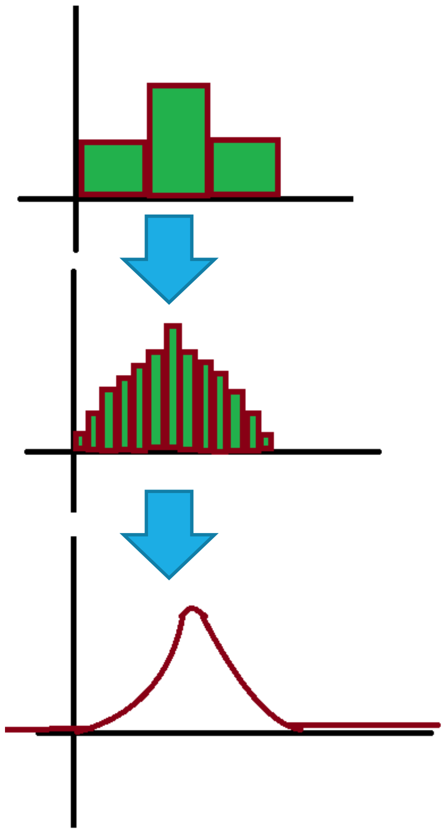
تعداد کل شیرهای چهار پرتاب سکه

$$P(X=0)=\frac{۱}{۱۶}$$
$$P(X=۱)=\frac{۴}{۱۶}$$
$$P(X=۲)=\frac{۶}{۱۶}$$
$$P(X=۳)=\frac{۴}{۱۶}$$
$$P(X=۴)=\frac{۱}{۱۶}$$

$$F(0)=P(0)=\frac{۱}{۱۶}$$
$$F(۱)=P(0)+P(۱)=\frac{۵}{۱۶}$$
$$F(۲)=P(0)+P(۱)+P(۲)=\frac{۱۱}{۱۶}$$
$$F(۳)=P(0)+P(۱)+P(۲)+P(۳)=\frac{۱۵}{۱۶}$$
$$F(۴)=P(0)+P(۱)+P(۲)+P(۳)+P(۴)=۱$$

$$F(x)=\begin{cases} 0, & x<0 \\ \frac{۱}{۱۶}, & 0\le x<۱ \\ \frac{۵}{۱۶}, & ۱\le x<۲ \\ \frac{۱۱}{۱۶}, & ۲\le x<۳ \\ \frac{۱۵}{۱۶}, & ۳\le x<۴ \\ ۱, & x\ge ۴ \end{cases}$$

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، ت۳ در همه‌جا مشتق‌پذیر است

«تچا» مشتق «ت۳» است (شبیه تفاضل در حالت گسسته)

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

$$P(x \le X \le x + \delta x) = f_X(x)\delta x$$

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

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

تابع چگالی احتمال – ویژگی‌ها

$$f_X(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f_X(x)dx = ۱$$

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

مثال

$f(x) = \begin{cases} ke^{-۳x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$ در چه صورتی چگالی احتمال است؟

$\int_{-\infty}^{+\infty} f_X(x)dx = ۱$

$\int_{-\infty}^{+\infty} ke^{-۳x}dx = \int_{0}^{+\infty} ke^{-۳x}dx = \lim_{t\to\infty} \frac{ke^{-۳x}}{-۳}\Big|_0^t = \frac{k}{۳} = ۱ \Longrightarrow k = ۳$

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

- تابع توزیع تجمیعی

تابع چگالی احتمال –

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

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

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

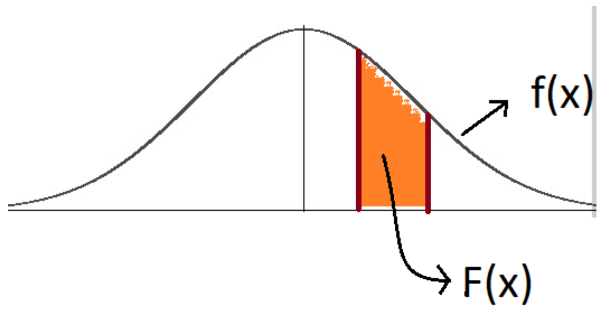
و این مشتق موجود است

مثال

$F(x) = \begin{cases} 0, & x \le 0 \\ x, & 0 < x < 1 \\ 1, & x \ge 1 \end{cases}$ تابع چگالی احتمال را بدست آورید و رسم کنید.

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 صرفا 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

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

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

X	Y
0	0
0	۱
۱	0
۱	۱
0	۲
۲	0

$$P(X=1,Y=0)=\frac{\binom{۳}{۱}\binom{۲}{0}\binom{۴}{۱}}{\binom{۹}{۲}}=\frac{۱}{۳}$$

$$P(X=1,Y=1)=\frac{\binom{۳}{۱}\binom{۲}{۱}\binom{۴}{0}}{\binom{۹}{۲}}=\frac{۱}{۶}$$

$$P(X=x,Y=y)=\frac{\binom{۳}{x}\binom{۲}{y}\binom{۴}{۲-x-y}}{\binom{۹}{۲}}, x,y\in\{0,۱,۲\}, 0\leq x+y\leq ۲$$

	$X=0$	$X=۱$	$X=۲$
$Y=0$	$\frac{۱}{۶}$	$\frac{۱}{۳}$	$\frac{۱}{۱۲}$
$Y=۱$	$\frac{۲}{۹}$	$\frac{۱}{۶}$	
$Y=۲$	$\frac{۱}{۳۶}$		

ویژگی‌ها

$$f(x, y) \geq 0$$

$$\sum_{x} \sum_{y} f(x, y) = 1$$

مثال $f(x,y) = kxy, \quad x,y =$ ۱و۲و۳

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

	$X=0$	$X=۱$	$X=۲$
$Y=0$	$\frac{۱}{۶}$	$\frac{۱}{۳}$	$\frac{۱}{۱۲}$
$Y=۱$	$\frac{۲}{۹}$	$\frac{۱}{۶}$	
$Y=۲$	$\frac{۱}{۳۶}$		

حالت پیوسته

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

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

ویژگی‌ها

$$f(x,y) \geq 0$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)dxdy = ۱$$

تابع توزیع توام

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

چگالی توام:

$$f(x,y) = \frac{\partial^{۲}}{\partial x \partial y} F(x,y)$$

مثال

چگالی توام X و Y

$$f(x,y)=\begin{cases} x+y, 0<x<1, 0<y<1 \\ 0 \quad \text{سایرجاها} \end{cases}$$

$$F(x,y)=\begin{cases} 0 & x\le 0 \text{ یا } y\le 0, \\ \frac{1}{2}xy(x+y), & 0<x<1, 0<y<1 \\ \frac{1}{2}y(1+y), & x\ge 1, 0<y<1 \\ \frac{1}{2}x(x+1), & 0<x<1, y\ge 1 \\ 1, & x\ge 1, y\ge 1 \end{cases}$$

توزیع‌های حاشیه‌ای- مثال

X	Y
0	0
0	۱
۱	0
۱	۱
0	۲
۲	0

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

توزیع احتمال X و توزیع احتمال Y

	$X = 0$	$X = 1$	$X = 2$	
$Y = 0$	$\frac{۱}{۶}$	$\frac{۱}{۳}$	$\frac{۱}{۱۲}$	$\frac{۷}{۱۲}$
$Y = 1$	$\frac{۲}{۹}$	$\frac{۱}{۶}$		$\frac{۷}{۱۸}$
$Y = 2$	$\frac{۱}{۳۶}$			$\frac{۱}{۳۶}$
	$\frac{۵}{۱۲}$	$\frac{۱}{۲}$	$\frac{۱}{۱۲}$	

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

$$f_{XY}(x,y)=\begin{cases}\frac{۲}{۳}(x+۲y), 0<x<۱, 0<y<۱\\ 0, & \text{درغیراینصورت}\end{cases}$$

چگالی‌های حاشیه‌ای X و Y

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^1 \frac{۲}{۳}(x+۲y)dy=\frac{۲}{۳}(x+۱)$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^1 \frac{۲}{۳}(x+۲y)dx=\frac{۱}{۳}(۱+۴y)$$

توزیع‌های توام و حاشیه‌ای – تابع چگالی

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

چگالی احتمال حاشیه‌ای X

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

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

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

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

چند متغیری گسسته

$$f(x_۱, x_۲, \ldots, x_n) = P(X_۱ = x_۱, X_۲ = x_۲, \ldots, X_n = x_n)$$

$$F(x_۱, x_۲, \ldots, x_n) = P(X_۱ \leq x_۱, X_۲ \leq x_۲, \ldots, X_n \leq x_n)$$

مثال

$f(x,y,z) = \frac{(x+y)z}{۶۳}, x = ۱,۲; y = ۱,۲,۳; z = ۱,۲$

$P(X = ۲, Y + Z \leq ۳)$؟

$$P(X = ۲, Y + Z \leq ۳) = f(۲,۱,۱) + f(۲,۱,۲) + f(۲,۲,۱)$$
$$= \frac{۳}{۶۳} + \frac{۶}{۶۳} + \frac{۴}{۶۳} = \frac{۱۳}{۶۳}$$

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

	$X=0$	$X=1$	$X=2$	
$Y=0$	$\frac{۱}{۶}$	$\frac{۱}{۳}$	$\frac{۱}{۱۲}$	$\frac{۷}{۱۲}$
$Y=1$	$\frac{۲}{۹}$	$\frac{۱}{۶}$		$\frac{۷}{۱۸}$
$Y=2$	$\frac{۱}{۳۶}$			$\frac{۱}{۳۶}$
	$\frac{۵}{۱۲}$	$\frac{۱}{۲}$	$\frac{۱}{۱۲}$	

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

$$P(X=0|Y=۱)=\frac{\frac{۲}{۹}}{\frac{۷}{۱۸}}=\frac{۴}{۷}$$

$$P(X=۱|Y=۱)=\frac{\frac{۱}{۶}}{\frac{۷}{۱۸}}=\frac{۳}{۷}$$

$$P(X=۲|Y=۱)=\frac{0}{\frac{۷}{۱۸}}=0$$

مثال

چگالی شرطی X با داشتن $Y = y$: $P\left(x \leq \frac{۱}{۲} | y = \frac{۱}{۲}\right)$

$$f_{XY}(x,y) = \begin{cases} \frac{۲}{۳}(x + ۲y), 0 < x < ۱, 0 < y < ۱ \\ 0, \quad \text{درغیراینصورت} \end{cases}$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_0^1 \frac{۲}{۳}(x + ۲y)dy = \frac{۲}{۳}(x + ۱)$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_0^1 \frac{۲}{۳}(x + ۲y)dx = \frac{۱}{۳}(۱ + ۴y)$$

مثال

$$f_{XY}(x,y)=\begin{cases}\frac{۲}{۳}(x+۲y), 0<x<۱, 0<y<۱\\ 0, \quad \text{درغیراینصورت}\end{cases}$$

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^1 \frac{۲}{۳}(x+۲y)dy=\frac{۲}{۳}(x+۱)$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^1 \frac{۲}{۳}(x+۲y)dx=\frac{۱}{۳}(۱+۴y)$$

چگالی شرطی X با داشتن $Y=y$: $P\left(x\le\frac{۱}{۲}|y=\frac{۱}{۲}\right)$

$$f(x|y)=\frac{f(x,y)}{f(y)}=\frac{\frac{۲}{۳}(x+۲y)}{\frac{۱}{۳}(۱+۴y)}=\begin{cases}\frac{(۲x+۴y)}{(۱+۴y)}, 0<x<۱\\ 0, \quad \text{درغیراینصورت}\end{cases}$$

$$f\left(x|y=\frac{۱}{۲}\right)=\frac{(۲x+۴y)}{(۱+۴y)}=\frac{\left(۲x+۴\times\frac{۱}{۲}\right)}{\left(۱+۴\times\frac{۱}{۲}\right)}=\frac{۲x+۲}{۳}$$

$$P\left(x\le\frac{۱}{۲}|y=\frac{۱}{۲}\right)=\int_0^{\frac{۱}{۲}}\frac{۲x+۲}{۳}dx=\frac{۵}{۱۲}$$

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 زیرمجموعهء A و B از اعداد حقیقی خواهیم داشت:

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

یا

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$$

همچنین

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

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

اثبات-

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

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

قاعدهٔ زنجیری

متغیرهای $X_1, \dots, X_k$

$$P_{X_1,\dots,X_k}(x_۱ \cap x_۲ \cap \cdots \cap x_k)$$

$$= P_{X_1}(x_۱) P_{X_۲|X_۱}(x_۲|x_۱) \dots P_{X_k|X_1,\dots,X_{k-1}}(x_k|x_۱ \cap x_۲ \cap \cdots \cap x_{k-۱})$$

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

$$f(x)=\frac{\binom{۲}{x}\binom{۱۰}{۳-x}}{\binom{۱۲}{۳}},\qquad x=0,۱,۲$$

x	0	۱	۲
$f(x)$	$\frac{۶}{۱۱}$	$\frac{۹}{۲۲}$	$\frac{۱}{۲۲}$

$$E(X)=0.\frac{۶}{۱۱}+۱.\frac{۹}{۲۲}+۲.\frac{۱}{۲۲}$$

مثال

$$f(x) = \begin{cases} \dfrac{۴}{\pi(۱ + x^{۲})}, 0 < x < ۱ \\ 0 \qquad\qquad\qquad \text{سایرجاها} \end{cases}$$

$$E(X) = \int_0^{۱} x \frac{۴}{\pi(۱ + x^{۲})} = \frac{۴}{\pi} \int_0^{۱} \frac{x}{۱ + x^{۲}}$$

امید ریاضی

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

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

$g(X) = ۲X^۲ + ۱$

$$\mathbb{E}[g(X)] = \sum_{x=۱}^{۶} (۲x^۲ + ۱).\frac{۱}{۶}$$
$$= (۲ \times ۱^۲ + ۱).\frac{۱}{۶} + \cdots + (۲ \times ۶^۲ + ۱).\frac{۱}{۶} = \frac{۹۴}{۳}$$

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

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

$$\mathbb{E}[a_۱X_۱ + a_۲X_۲ + \cdots + a_nX_n] = a_۱\mathbb{E}[X_۱] + a_۲\mathbb{E}[X_۲] + \cdots + a_n\mathbb{E}[X_n]$$

متغیر تصادفی گسستهٔ X

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

مثال-

احتمال جمع امید ریاضی حاصل از ریختن سه تاس

- $X = X_۱ + X_۲ + X_۳$

$$E[X] = E[X_۱] + E[X_۲] + E[X_۳] = ۳\left(\frac{۷}{۲}\right) = \frac{۲۱}{۲}$$

مثال-

امید ریاضی مت دوجمله‌ای با پارامترهای n و p

- $X = X_۱ + X_۲ + \cdots + X_n$

$$X_i = \begin{cases} ۱ \\ 0 \end{cases}$$

متغیر تصادفی برابر ۱ در صورت پیروزی آزمایش i-ام در غیر این‌صورت برابر 0

- X_i مت برنولی
  - پس $\mathrm{E}[X_i] = ۱p + 0(۱ - p) = p$

$$E[X] = E[X_۱] + E[X_۲] + \cdots + E[X_n] = np$$

وردائی یا واریانس

سنجش پراکندگی

چگونگی تمرکز مت X حول میانگین

$$
\begin{aligned}
Var[X] &= E\left[(X - E[X])^{۲}\right] \\
&= E\left[X^{۲}\right] - E[X]^{۲}
\end{aligned}
$$

واردائی یا واریانس – ویژگی‌ها

$$Var[a] = 0$$

$$Var[af(X)] = a^{۲} Var[f(X)]$$

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

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

چند توزیع پراستفاده

بررسی میانگین و وردائی

توزیع‌های احتمال گسسته

چگالی‌های احتمال پیوسته

توزیع یکنواخت

برامد تمام خروجی‌ها دارای احتمال برابر

$$f(x) = \frac{1}{k}, x = x_۱, \dots, x_k$$

توزیع برنولی

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

$X \sim$ برنولی(p)

$0 \leq p \leq 1$

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

شیر=1

خط=0

توزیع برنولی

$X \sim$ برنولی(p)

$0 \leq p \leq 1$

شیر=1

خط=0

$$E[X] = p$$
$$Var[X] = p(1 - p)$$

توزیع هندسی

x آزمایش

- با دو حالت شکست و پیروزی در هر آزمایش
- احتمال پیروزی p احتمال شکست q
  - $p + q = ۱$

احتمال پیروزی در آزمایش x-اُم (شکست $x - ۱$ مرحلهٔ قبلی)

- توزیعی گسسته

$$g(x;p) = (۱ - p)^{x-۱}p$$

- میانگین

$$\mu = \frac{۱}{p}$$

احتمال پیروزی در آزمایش $x + ۱$-اُم (شکست x مرحلهٔ قبلی)

- توزیعی گسسته

$$g(x;p) = (۱ - p)^{x}p$$

- میانگین

$$\mu = \frac{۱ - p}{p}$$

توزیع دوجمله‌ای

امتحان‌های تکراری

احتمال تمام پیروزی‌ها برابر با p

هر پرتاب (آزمایش) مستقل از دیگر آزمایش‌ها

$X \sim \text{دوجمله‌ای}(x; n, p)$

$0 \leq p \leq ۱$

$$b(x; n, p) = \binom{n}{x} p^x (۱ - p)^{n-x}, \qquad x = 0, ۱, \ldots, n$$

مثال

بهبود ۷ نفر از ۱۰ نفر با احتمال بهبود ۰.۸

$$b(۷;۱۰,۰.۸)=\binom{۱۰}{۷}۰.۸^{۷}(۱-۰.۸)^{۳}\approx 0.۲$$

ویژگی‌ها

احتمال تجمعی

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

$$b(x;n,p) = b(n-x;n,1-p)$$

$$\mu = np$$

$$\sigma^{۲} = np(1-p)$$

توزیع فوق‌هندسی

نمونه‌گیری بدون جایگذاری

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

$$X \sim \text{فوق‌هندسی}(x; n, N, M)$$

$$h(x; n, N, M) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \qquad x = 0, ۱, ۲, ۳, \dots; x \leq M; n - x \leq N - M$$

ویژگی‌ها

$$p = \frac{M}{N} \Longrightarrow \mu = \frac{nM}{N} = np$$

$$p = \frac{M}{N} \Longrightarrow \sigma^۲ = \frac{nM(N-M)(N-n)}{N^۲(N-1)} = \frac{nM(N-M)(N-n)}{N^۲(N-1)} = \mathrm{n}\frac{\mathrm{M}}{\mathrm{N}}\frac{\mathrm{N}\left(۱-\frac{M}{N}\right)}{N}\frac{\mathrm{N}\left(۱-\frac{n}{N}\right)}{N\left(۱-\frac{۱}{N}\right)} = np(۱-p)$$

با شرط N بزرگ و N $n \ll N$

- عدم‌تفاوت بین نمونه‌گیری با جاگذاری و بدون جای‌گذاری
- امکان استفاده از تقریب احتمال فوق هندسی $h(x; n,N,M)$ با معادله توزیع دوجمله‌ای $b\left(x; n, \frac{M}{N}\right)$
- حکم برقرار است

توزیع پواسن

سختی محاسبهٔ 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.

$$p(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!}, x=0, ۱, ۲, \ldots$$

توزیع چندجمله‌ای

$$f(x_۱, x_۲, \dots, x_k; n, p_۱, p_۲, \dots, p_k) = \binom{n}{x_۱, x_۲, \dots, x_k} p_۱^{x_۱} p_۲^{x_۲} \dots p_k^{x_k}$$

$$x_۱ + x_۲ + \cdots + x_k = n$$

$$p_۱ + p_۲ + \cdots + p_k = 1$$

توزیع فوق‌هندسی چندمتغیره

$$f(x_۱, x_۲, \dots, x_k; n, M_۱, M_۲, \dots, M_k) = \frac{\binom{M_۱}{x_۱}\binom{M_۲}{x_۲}\cdots\binom{M_k}{x_k}}{\binom{N}{n}}$$

$$x_i = 0, ۱, ۲, \dots, n$$

$$x_i \le M_i$$

$$x_۱ + x_۲ + \cdots + x_k = n$$

$$M_۱ + M_۲ + \cdots + M_k = N$$

توزیع‌های پیوسته

توزیع یکنواخت

$X \sim \text{یکنواخت}(a, b)$

$a < b$

چگالی احتمال یکسان برای هر نقطه بین a و b

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

چگالی احتمال نمائی

زمان انتظار جهت وقوع پیشامد بعدی با میانگین $\frac{۱}{\lambda}$

$\theta > 0$

چگالی احتمال کاهشی روی مقادیر حقیقی نامنفی

$$e(x;\theta) = \begin{cases} \frac{۱}{\theta} e^{-\frac{x}{\theta}}, x > 0 \\ 0 \text{سایرجاها} \end{cases}$$

یا

$$e(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, x > 0 \\ 0 \text{سایرجاها} \end{cases}$$

$$F(x;\lambda) = \begin{cases} ۱ - e^{-\lambda x}, > 0 \\ 0 \text{سایرجاها} \end{cases}$$

چگالی احتمال نمائی

مثال‌ها

- زمان تلفن نفر بعدی
- زمان بین چاپ دو مقاله
- فاصلهٔ زمانی بین دو ارباب رجوع به اداره پست
- فاصله زمانی دو تصادفی در دوربرگردان بعد میدان کوهنورد جاده گاوازنگ
- دو زمین‌لرزه در ایران
- زمان بین دو بارندگی

$$E[X] = \frac{1}{\lambda}$$

$$Var[X] = \frac{۱}{\lambda^{۲}}$$

چگالی احتمال نمائی

مثال- به طور متوسط هر سه ساعت یکبار تصادفی در جاده گاوازنگ رخ می‌دهد. احتمال تصادفی بعدی در بازه سه و هفت ساعت چقدر است؟

$$E[X] = \frac{1}{\lambda} = ۳$$

$$P(3 < X < 7) = F(۷) - F(۳) = \left(1 - e^{-\frac{7}{3}}\right) - (1 - e^{-1}) \approx 0.27$$

چگالی احتمال نمائی

خاصیت بی‌حافظگی

توزیع نمایی و توزیع هندسی

- هر دو بی‌حافظه
- اگر X مت نمایی آن‌گاه جز صحیح X دارای توزیع هندسی

چگالی احتمال گاما

زمان انتظار تا n-امین وقوع پیشامد با میانگین ۱/λ

$$g(x;n,\lambda) = \begin{cases} \dfrac{\lambda(\lambda x)^{n-۱}e^{-\lambda x}}{(n-۱)!}, x \geq 0 \\ 0, x < 0 \end{cases}$$

چگالی احتمال گاما

تعمیم به α

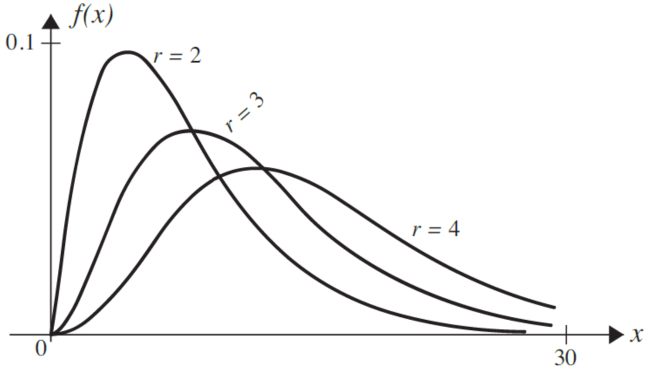
$$g(x;\alpha,\theta)=\begin{cases}\dfrac{۱}{\theta^{\alpha}\Gamma(\alpha)}x^{\alpha-۱}e^{-\frac{x}{\theta}}, x\geq 0\\ 0, x<0\end{cases}$$

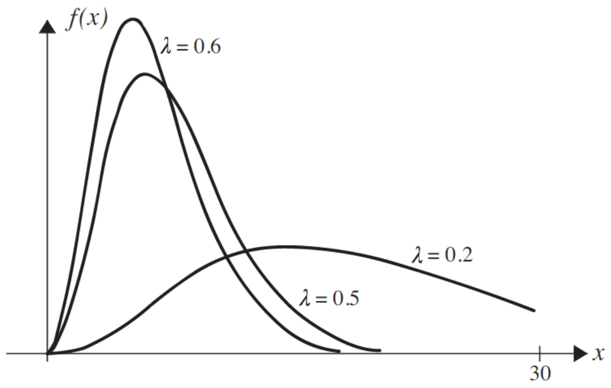
یا

$$g(x;\alpha,\lambda)=\begin{cases}\dfrac{\lambda(\lambda x)^{\alpha-۱}e^{-\lambda x}}{\Gamma(\alpha)}, x\geq 0\\ 0, x<0\end{cases}$$

$$\Gamma(\alpha)=\int_{0}^{\infty}e^{-x}x^{\alpha-۱}dx$$

$$\Gamma(n)=(\mathrm{n}-۱)!$$





چگالی احتمال گاما

مثال- فرض در هر ثانیه ماده رادیواکتیوی چهار ذره منتشر می‌کند. احتمال انتشار دو ذره حداقل دو ثانیه وقت لازم داشته باشد را محاسبه کنید.

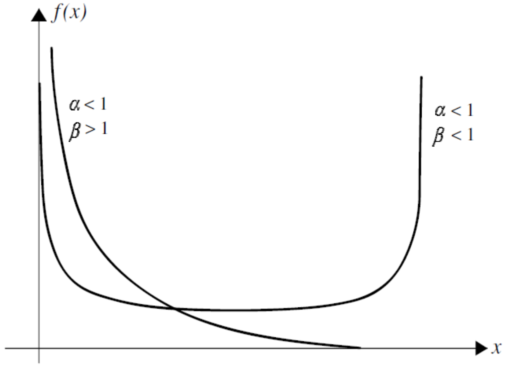
$\lambda = ۴$

$n = ۲$

x زمان تا انتشار دومین ذره. توزیع گاما $(۲,۴)$

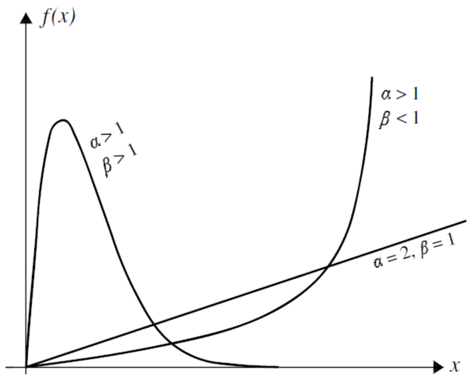
$$P(X \geq ۲) = \int_{۲}^{\infty} \frac{4e^{-4x}(4x)^{2-1}}{(2-1)!} dx = \int_{۲}^{\infty} 16e^{-4x} dx \approx= 0.003$$

چگالی احتمال بتا



$$g(x;\alpha,\beta)=\begin{cases}\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, 0<x<1\\ 0\text{سایرجاها}\end{cases}$$

$$\alpha,\beta>0$$



مناسب برای مت با تلاطم بین دو کران پایین و بالا

مثال

- نسبت افراد جامعه گوش دهنده به علیرضا قربانی و یا غلامحسین بنان در فاصله زمانی معین
- در صد سطح مزارع زیر کشت گندم

چگالی احتمال بتا

مثال- میله‌ای به طول l داریم. در نقطه‌ای تصادفی ضربه‌ای به آن وارد می‌شود و موجب شکستن آن در فاصلهٔ X از مبدا میله می‌شود. اگر X/l داری توزیع بتا باشد و $\alpha = \beta = 3$ احتمال $P(l/5 < X < l/4)$؟

$$P\left(\frac{l}{5} < X < \frac{l}{4}\right) = P\left(\frac{1}{5} < \frac{X}{l} < \frac{1}{4}\right) = \int_{\frac{1}{5}}^{\frac{1}{4}} \frac{\Gamma(6)}{\Gamma(3)\Gamma(3)} x^{3-1}(1-x)^{3-۱} dx$$

$$= ۳۰ \int_{\frac{1}{5}}^{\frac{1}{4}} x^{۲}(1-x)^2 dx = 0.046$$

چگالی احتمال نرمال (گاوسی یا زنگوله‌ای)

$$n(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{۲\pi}}e^{-\frac{1}{۲}(\frac{x-\mu}{\sigma})^{۲}},-\infty<x<+\infty,\sigma>0$$

توزیع نرمال استانده

- $\mu=0,\sigma=۱$

اساس نظریهٔ آمار

توزیع ها

نوع	فرایند	ضرائب	چگالی یا جرم	تابع تجمیع	تم‌گ	میانگین	وردائی	چولگی
گسسته	برنولی	θ	$f(x;\theta)=\theta^{x}(۱-\theta)^{۱-x}, x=0,۱$	$F(x;\ \theta)=\begin{cases}0, & x<0\\ ۱-\theta, 0\le x\le ۱\\ ۱, & x\ge ۱\end{cases}$	$۱-\theta+\theta e^{t}$	θ	$\theta(1-\theta)$	$\frac{۱-۲\theta}{\sqrt{\theta(۱-\theta)}}$
	هندسی	θ	$f(x;\theta)=(۱-\theta)^{x-۱}\theta$	$F(x;\ \theta)=۱-(۱-\theta)^{x}$	$\frac{\theta e^{t}}{۱-(۱-\theta)e^{t}}$	$\frac{۱}{\theta}$	$\frac{۱-\theta}{\theta^{۲}}$	$\frac{۲-\theta}{\sqrt{(۱-\theta)}}$
	دوجمله‌ای	n,θ	$f(x;n,\theta)=\binom{n}{x}\theta^{x}(۱-\theta)^{n-x}$	$F(x;n,\theta)=\sum_{i=0}^{x}\binom{n}{i}\theta^{i}(۱-\theta)^{n-i}$	$(\theta e^{t}+۱-\theta)^{n}$	$n\theta$	$n\theta(۱-\theta)$	$\frac{۱-۲\theta}{\sqrt{n\theta(۱-\theta)}}$
	فوق‌هندسی	n,N,M	$f(x;n,N,M)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	$F(x;n,N,M)=\sum_{i=0}^{x}\frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$		$n\frac{M}{N}$	$n\frac{M}{N}\frac{(N-M)}{N}\frac{(N-n)}{N-۱}$	
	پواسن	λ	$f(x,\lambda)=\frac{\lambda^{x}e^{-\lambda}}{x!}$	$F(x,\lambda)=e^{-\lambda}\sum_{i=0}^{x}\frac{\lambda^{i}}{i!}$	$e^{\lambda(e^{t}-۱)}$	λ	λ	$\frac{۱}{\sqrt{\lambda}}$
پیوسته	نمائی	λ	$f(x,\lambda)=\lambda e^{-\lambda x}$	$F(x,\lambda)=۱-e^{-\lambda x}$	$\frac{\lambda}{\lambda-t}$	$\frac{۱}{\lambda}$	$\frac{۱}{\lambda^{۲}}$	۲
	گاما	α,λ	$f(x;\alpha,\lambda)=\begin{cases}\frac{\lambda(\lambda x)^{\alpha-۱}e^{-\lambda x}}{\Gamma(\alpha)}, x\ge 0\\ 0, x<0\end{cases}$		$\left(\frac{\lambda}{\lambda-t}\right)^{\alpha}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^{۲}}$	$\frac{۲}{\sqrt{\alpha}}$
	بتا	α,β	$f(x;\alpha,\beta)=\begin{cases}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-۱}, 0<x<1\\ \text{سایر جاها}\ 0\end{cases}$			$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^{۲}(\alpha+\beta+۱)}$	
	گاوس	$\mu,\sigma^{۲}$	$f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{۲\pi}}e^{-\frac{1}{۲}\left(\frac{x-\mu}{\sigma}\right)^{۲}}$		$e^{\frac{\sigma^{۲}t^{۲}}{۲}+\mu t}$	μ	$\sigma^{۲}$	0

امید ریاضی

گسسته

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

پیوسته

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

$E[XY] = E[X]E[Y]$؟

امید ریاضی

$g: \mathbb{R}^۲ \rightarrow \mathbb{R}$

گسسته

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

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 (هم‌وردائی)

استفاده از امید ریاضی جهت مطالعه روابط بین دو مت

$$Cov[X,Y] = E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] = E[XY] - E[X]E[Y]$$

$Cov[X,Y] = 0$؟

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس

ویژگی‌ها

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

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

$$Var[X + Y] = Var[X] + Var[Y] + ۲Cov[X,Y]$$

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	Y
0	0
0	۱
۱	0
۱	۱
0	۲
۲	0

	$X=0$	$X=1$	$X=2$	
$Y=0$	$\frac{۱}{۶}$	$\frac{۱}{۳}$	$\frac{۱}{۱۲}$	$\frac{۷}{۱۲}$
$Y=1$	$\frac{۲}{۹}$	$\frac{۱}{۶}$		$\frac{۷}{۱۸}$
$Y=2$	$\frac{۱}{۳۶}$			$\frac{۱}{۳۶}$
	$\frac{۵}{۱۲}$	$\frac{۱}{۲}$	$\frac{۱}{۱۲}$	

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

توزیع احتمال X و توزیع احتمال Y

$$\mu_{XY} = E(XY) = 0 \times 0 \times \frac{۱}{۶} + 0 \times ۱ \times \frac{۲}{۹} + 0 \times ۲ \times \frac{۱}{۳۶} + ۱ \times 0 \times \frac{۱}{۳} + ۱ \times ۱ \times \frac{۱}{۶} + ۲ \times 0 \times \frac{۱}{۱۲} = \frac{۱}{۶}$$

$$\mu_X = E(X) = 0 \times \frac{۵}{۱۲} + ۱ \times \frac{۱}{۲} + ۲ \times \frac{۱}{۱۲} = \frac{۲}{۳}$$

$$\mu_Y = E(Y) = 0 \times \frac{۷}{۱۲} + ۱ \times \frac{۷}{۱۸} + ۲ \times \frac{۱}{۳۶} = \frac{۴}{۹}$$

$$\sigma_{XY} = \frac{۱}{۶} - \frac{۲}{۳} \times \frac{۴}{۹} = -\frac{۷}{۵۴}?$$

تابع چگالی توأم X و Y

$$f(x,y)=\frac{۱}{y}e^{-\left(y+\frac{x}{y}\right)},0<x,y<\infty$$

الف- تابع چگالی است؟

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)dydx=\int_{0}^{\infty}\int_{0}^{\infty}\frac{۱}{y}e^{-\left(y+\frac{x}{y}\right)}dydx$$

$$=\int_{0}^{\infty}e^{-y}\int_{0}^{\infty}\frac{۱}{y}e^{-\frac{x}{y}}dxdy=\int_{0}^{\infty}e^{-y}dy=۱$$

تابع چگالی توأم X و Y

$$f(x,y)=\frac{١}{y}e^{-\left(y+\frac{x}{y}\right)},0<x,y<\infty$$

ب- هم‌وردائی؟

$$f_Y(y)=e^{-y}\int_0^\infty \frac{١}{y}e^{-\frac{x}{y}}dx=e^{-y}\Longrightarrow E[Y]=\int_0^\infty ye^{-y}dy=١$$

$$E[X]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dydx=\int_0^\infty e^{-y}\int_0^\infty \frac{x}{y}e^{-\frac{x}{y}}dxdy=\int_0^\infty ye^{-y}dy=١$$

$$E[XY]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dydx=\int_0^\infty ye^{-y}\int_0^\infty \frac{x}{y}e^{-\frac{x}{y}}dxdy=\int_0^\infty y^{٢}e^{-y}dy=٢$$

$$Cov[X,Y]=E[XY]-E[X]E[Y]=١$$

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس(هم‌وردائی)

ویژگی‌ها

$$Cov[aX+b, cY+d] = acCov(X,Y)$$

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

$$Cov[X, Y+Z] = Cov[Y,X] + Cov(X,Z) \Rightarrow Cov\left(\sum_{i=۱}^{n} X_i, \sum_{j=۱}^{m} Y_j\right) = \sum_{i=۱}^{n}\sum_{j=۱}^{m} Cov(X_i, Y_j)$$

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

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس(هم‌وردائی)

$$Var\left(\sum_{i=۱}^{n} X_i\right)$$

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس(هم‌وردائی)

$$Var\left(\sum_{i=۱}^{n} X_i\right) = Cov\left(\sum_{i=۱}^{n} X_i, \sum_{j=۱}^{n} X_j\right) = \sum_{i=۱}^{n}\sum_{j=۱}^{n} Cov(X_i, X_j)$$

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس(هم‌وردائی)

$$Var\left(\sum_{i=۱}^{n} X_i\right) = Cov\left(\sum_{i=۱}^{n} X_i\,, \sum_{j=۱}^{n} X_j\right) = \sum_{i=۱}^{n}\sum_{j=۱}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=۱}^{n} Cov(X_i, X_i) + \sum_{i=۱}^{n}\sum_{j\neq i} Cov(X_i, X_j)$$

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس(هم‌وردائی)

$$Var\left(\sum_{i=۱}^{n} X_i\right) = Cov\left(\sum_{i=۱}^{n} X_i , \sum_{j=۱}^{n} X_j\right) = \sum_{i=۱}^{n}\sum_{j=۱}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=۱}^{n} Cov(X_i, X_i) + \sum_{i=۱}^{n}\sum_{j\neq i} Cov(X_i, X_j)$$

$$Var\left(\sum_{i=۱}^{n} X_i\right) = \sum_{i=۱}^{n} Var(X_i) + ۲\sum_{i=۱}^{n}\sum_{j<i} Cov(X_i, X_j)$$

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس(هم‌وردائی)

$$Var\left(\sum_{i=۱}^{n} X_i\right) = Cov\left(\sum_{i=۱}^{n} X_i , \sum_{j=۱}^{n} X_j\right) = \sum_{i=۱}^{n}\sum_{j=۱}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=۱}^{n} Cov(X_i, X_i) + \sum_{i=۱}^{n}\sum_{j\neq i} Cov(X_i, X_j)$$

$$Var\left(\sum_{i=۱}^{n} X_i\right) = \sum_{i=۱}^{n} Var(X_i) + ۲\sum_{i=۱}^{n}\sum_{j<i} Cov(X_i, X_j)$$

در صورت استقلال X_i-ها از یکدیگر

$$Var\left(\sum_{i=۱}^{n} X_i\right) = \sum_{i=۱}^{n} Var(X_i)$$

توزیع‌های توأم و حاشیه‌ای – امید ریاضی و کوواریانس

ویژگی‌ها

$$Var[X+Y] = Var[X] + Var[Y] + ۲Cov[X,Y]$$

استقلال دو متغیر

$$Var[X+Y] = Var[X] + Var[Y]$$

$$Var[aX+bY] = a^۲Var[X] + b^۲Var[Y] + ۲abCov(X,Y)$$

مثال

سه متغیر تصادفی X و Y و Z

$$\mu_X = ۲, \mu_Y = -۳, \mu_Z = ۴$$

$$\sigma_X^۲ = ۱, \sigma_Y^۲ = ۵, \sigma_Z^۲ = ۲$$

$$Cov(X,Y) = -۲, Cov(X,Z) = -۱, Cov(Y,Z) = ۱$$

میانگین و وردائی $W = 3X - Y + 2Z$

$$E(W) = E(۳X - Y + ۲Z) = ۳E(X) - E(Y) + ۲E(Z) = ۳ \times ۲ - (-۳) + ۲ \times ۴ = ۱۷$$

$$Var(W) = ۹Var(X) + Var(Y) + ۴Var(Z) - ۶Cov(X,Y) + ۱۲Cov(X,Z) - ۴Cov(Y,Z)$$

$$= ۹ \times ۱ + ۵ + ۴ \times ۲ - ۶ \times (-۲) + ۱۲(-۱) - ۴ \times ۱ = ۱۸$$

توابع مولد گشتاور

تابع مولد گشتاور $\phi(t)$ مت X

$$\phi(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx} p(x), & \text{گسسته } X \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx, & \text{پیوسته} X \end{cases}$$

- دلیل نام
  - بدست آوردن تمامی گشتاورهای X با مشتق‌گیری متوالی از $\phi(t)$

توابع مولد گشتاور

تابع مولد گشتاور $\phi(t)$ مت X

$$\phi(t) = E[e^{tX}]$$

$$\phi'(t) = \frac{d}{dt}E[e^{tX}] = E\left[\frac{d}{dt}(e^{tX})\right] = E[Xe^{tX}] \Rightarrow \phi'(t) = E[Xe^{tX}]$$

$$\phi'(0) = E[X]$$

- به طریق مشابه

$$\phi''(t) = \frac{d}{dt}\phi'(t) = \frac{d}{dt}E[Xe^{tX}]) = E\left[\frac{d}{dt}Xe^{tX}\right] = E\left[X^{۲}e^{tX}\right]$$

$$\Rightarrow \phi''(0) = E\left[X^{۲}\right]$$

به طریق اولی، قضیه

$$\phi^{n}(0) = E[X^{n}], n \geq ۱$$

توابع مولد گشتاور

قضیه

$$\phi^n(0) = E[X^n], n \geq ۱$$

- نتیجه

$$\phi(t) = \sum_{i=0}^{\infty} \frac{E[X^i]}{i!} t^i = \sum_{i=0}^{\infty} \frac{\Phi^{(i)}(0)}{i!} t^i$$

مثال

گشتاور توزیع برنولی با پارامتر p را بدست آورید.

$$\phi(t) = E[e^{tx}] = (1-p)e^{t\times 0} + pe^{t\times 1} = 1 - p + pe^{t}$$

پس

$$\phi'(t) = pe^{t}$$
$$\phi'(0) = E[X] = p$$

$$\phi^{(n)}(0) = E[X^n] = p$$

مثال- توزیع دوجمله‌ای

توزیع دوجمله‌ای با پارامترهای n و p

$$\phi(t) = E[e^{tX}] = \sum_{i=0}^{n} e^{ti} \binom{n}{i} p^i (١ - p)^{n-i}$$

$$= \sum_{i=0}^{n} \binom{n}{i} (pe^t)^i (١ - p)^{n-i}$$

$$\phi(t) = (pe^t + ١ - p)^n$$

پس

$$\phi'(t) = n(pe^t + ١ - p)^{n-١} pe^t$$

$$\phi'(0) = E[X] = np$$

مثال

توزیع دوجمله‌ای با پارامترهای n و p

$$\phi(t) = (pe^t + ۱ - p)^n$$

پس

$$\phi'(t) = n(pe^t + ۱ - p)^{n-۱} pe^t$$
$$\phi'(0) = E[X] = np$$

$$\phi''(0) = n(n-۱)(pe^t + ۱ - p)^{n-۲}(pe^t)^۲ + n(pe^t + ۱ - p)^{n-۱} pe^t$$
$$\phi''(0) = E\left[X^۲\right] = n(n-۱)p^۲ + np$$

مثال- توزیع پواسن

توزیع پواسن با میانگین λ

$$\phi(t) = E[e^{tX}] = \sum_{i=0}^{\infty} \frac{e^{ti} e^{-\lambda} \lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda e^t)^i}{i!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - ۱)}$$

$$\phi(t) = e^{\lambda(e^t - ۱)}$$

پس

$$\phi'(t) = \lambda e^t e^{\lambda(e^t - ۱)}$$

$$\phi''(t) = (\lambda e^t)^۲ e^{\lambda(e^t - ۱)} + \lambda e^t e^{\lambda(e^t - ۱)}$$

در نتیجه

$$E[X] = \phi'(0) = \lambda$$

$$E[X^۲] = \phi''(0) = \lambda^۲ + \lambda$$

$$Var(X) = E[X^۲] - (E[X])^۲ = \lambda$$

مثال- توزیع نمایی

توزیع نمایی با پارامتر λ

$$\phi(t) = E[e^{tX}] = \frac{\lambda}{\lambda - t}, t < \lambda$$

پس

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^{۲}}$$

$$\phi''(t) = \frac{۲\lambda}{(\lambda - t)^{۳}}$$

در نتیجه

$$E[X] = \phi'(0) = \frac{۱}{\lambda}$$

$$E\left[X^{۲}\right] = \phi''(0) = \frac{۲}{\lambda^{۲}}$$

$$Var(X) = E\left[X^{۲}\right] - (E[X])^{۲} = \frac{۱}{\lambda^{۲}}$$

مثال- توزیع نرمال

توزیع نرمال با پارامترهای μ و $\sigma^۲$

$$\phi(t) = E[e^{tX}] = e^{\frac{\sigma^۲ t^۲}{۲} + \mu t}$$

پس

$$\phi'(t) = (\mu + t\sigma^۲) e^{\frac{\sigma^۲ t^۲}{۲} + \mu t}$$

$$\phi''(t) = (\mu + t\sigma^۲)^۲ e^{\frac{\sigma^۲ t^۲}{۲} + \mu t} + \sigma^۲ e^{\frac{\sigma^۲ t^۲}{۲} + \mu t}$$

در نتیجه

$$E[X] = \phi'(0) = \mu$$

$$E[X^۲] = \phi''(0) = \mu^۲ + \sigma^۲$$

$$Var(X) = E[X^۲] - (E[X])^۲ = \sigma^۲$$

ویژگی‌ها

$$\phi_X(t) = E[e^{tX}] = E(\cos tX) + iE(\sin tX)$$

قضیه

$$\phi_{X_۱+X_۲}(t) = \phi_{X_۱}(t)\phi_{X_۲}(t)$$

- تعمیم قضیهٔ اخیر؟

$$\phi_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} \phi_{X_i}(t)$$

قضیه

$$\forall t: \phi_{X_۱}(t) = \phi_{X_۲}(t) \Longrightarrow \forall x: F_{X_۱}(x) = F_{X_۲}(x)$$

عکس قضیه چطور؟

کران‌های احتمالی

گاهی اوقات عدم اطلاع کافی جهت محاسبه احتمال یا امید ریاضی

گاهی اوقات پیچیدگی حل مسئله

گاهی اوقات کفایت گزارش نتیجه‌ای عمومی و گسترش‌پذیر به طیف وسیعی از مسائل

- استفاده از کران‌های احتمالی به جای محاسبهٔ دقیق احتمال و یا کمیت‌های لازم.

کران اجتماع و تعمیم آن

کران اجتماع یا نامساوی بول

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &\leq P(A) + P(B). \end{aligned}$$

$$\begin{aligned} P(A \cup B \cup C) &= P\big((A \cup B) \cup C\big) \\ &\leq P(A \cup B) + P(C) \\ &\leq P(A) + P(B) + P(C). \end{aligned}$$

به ازای پیشامدهای $A_۱, A_۲, \ldots, A_n$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

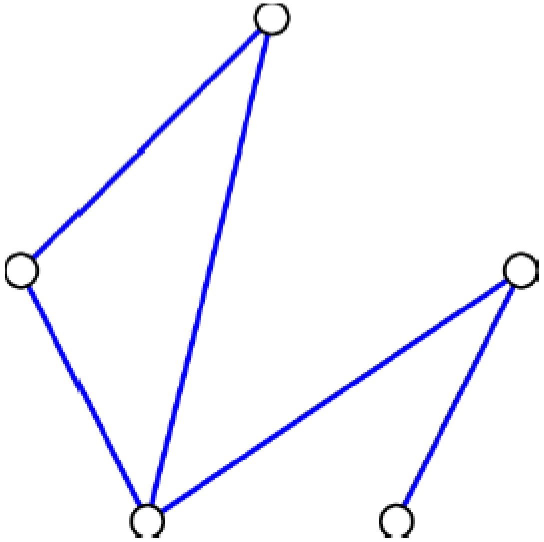
کران اجتماع و تعمیم آن

مثال- گراف‌های تصادفی

- مناسب در تحلیل شبکه‌های احتماعی، شبکه‌های بی‌سیم، اینترنت

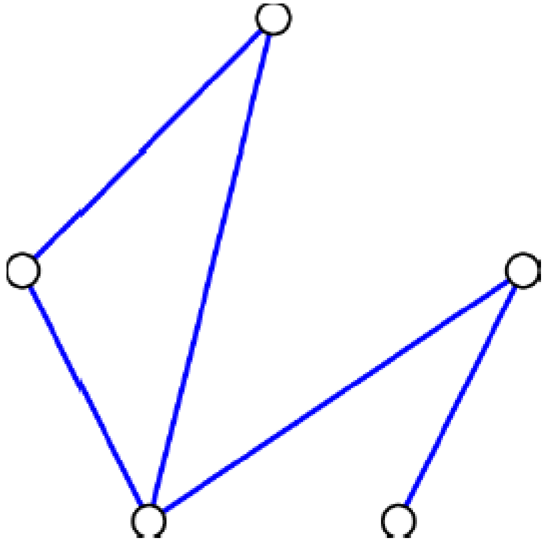
مدل اردوش-رنیی G(n,p) از مدل‌های ساده گراف‌های تصادفی

- n راس
- هر دو راس متصل با احتمال p
- استقلال وجود هر یال از دیگر یال‌های گراف
- نمونه G(5,1/2)



سوال- احتمال وجود راس ایزوله

کران اجتماع و تعمیم آن



فرض شبکه دارای **n** راس

فرض A_i به معنای پیشامد ایزوله بودن راس **i**–ام

پس احتمال ایزوله بودن حداقل یک راس برابر با

$$B_n = \bigcup_{i=1}^{n} A_i.$$

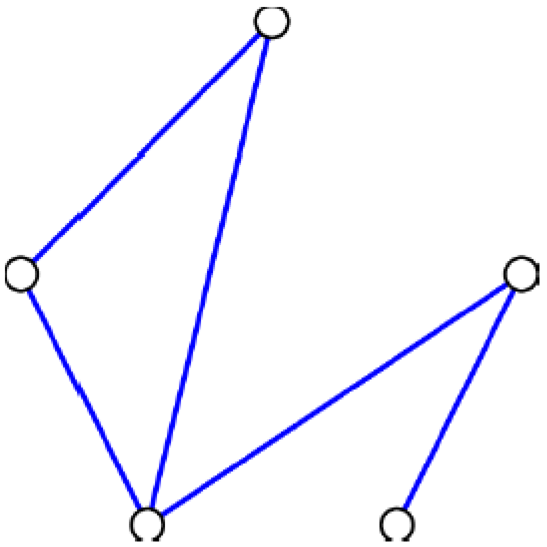
پس با استفاده از کران اجتماع داریم

$$P(B_n) = P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i).$$

به ازای هر **i** و **j** داریم: $P(A_i) = P(A_j)$، در نتیجه

$$P(B_n) \leq nP(A_1).$$

کران اجتماع و تعمیم آن



به ازای هر i و j داریم: $P(A_i) = P(A_j)$، در نتیجه $P(B_n) \leq nP(A_۱)$

پیشامد A_1 معادل با وصل نبود راس ۱ با هیچ از $n - ۱$ راس دیگر

- استقلال اتصالات از یکدیگر پس

$$P(A_1) = (1 - p)^{n-1}.$$

$$P(B_n) \leq n(1 - p)^{n-1}$$

کران اجتماع و تعمیم آن

به ازای پیشامدهای $A_1, A_2, \ldots, A_n$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i);$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j);$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k).$$

تمرین- اثبات کنید

$$P(B_n) \geq n(1-p)^{n-1} - \binom{n}{2}(1-p)^{2n-3}$$

نامساوی مارکوف

اگر X متغیر تصادفی باشد که صرفا مقادیر نامنفی را می‌پذیرد، آن‌گاه به ازای هر $a > 0$

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

نامساوی مارکوف

اگر X متغیر تصادفی باشد که صرفا مقادیر نامنفی را می‌پذیرد، آن‌گاه به ازای هر a>0

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

اثبات

$$E[X] = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \geq \int_a^\infty xf(x)dx \geq \int_a^\infty af(x)dx$$

$$= a\int_a^\infty f(x)dx = aP\{X \geq a\}$$

نامساوی مارکوف

اگر X متغیر تصادفی دوجمله‌ای(n,p) باشد. کران بالای $P(X \geq \alpha\mathrm{n})$؟ فرض p<$\alpha$<1

- $P = \frac{1}{2}, \alpha = \frac{3}{4}$

$$P\{X \geq \alpha n\} \leq \frac{E[X]}{\alpha n} = \frac{np}{\alpha n} = \frac{p}{\alpha n}$$

$$P\left\{X \geq \frac{3n}{4}\right\} \leq \frac{2}{3}$$

نامساوی چبیشف

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

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

یا

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

اثبات

$$\sigma^۲ = E\left[(X-\mu)^۲\right]$$

$$\sigma^۲ = \int_{-\infty}^{+\infty} (x-\mu)^۲ f(x)dx$$

$$\sigma^۲ = \int_{-\infty}^{\mu-k\sigma} (x-\mu)^۲ f(x)dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (x-\mu)^۲ f(x)dx + \int_{\mu+k\sigma}^{+\infty} (x-\mu)^۲ f(x)dx$$

انتگرل $(x-\mu)^۲ f(x)$ نامنفی $\Leftarrow$ پس

$$\sigma^۲ \geq \int_{-\infty}^{\mu-k\sigma} (x-\mu)^۲ f(x)dx + \int_{\mu+k\sigma}^{+\infty} (x-\mu)^۲ f(x)dx$$

چون $x \leq \mu - k\sigma$ یا $x \geq \mu + k\sigma$ پس $(x-\mu)^۲ \geq k^۲\sigma^۲$. در نتیجه

$$\sigma^۲ \geq \int_{-\infty}^{\mu-k\sigma} k^۲\sigma^۲ f(x)dx + \int_{\mu+k\sigma}^{+\infty} k^۲\sigma^۲ f(x)dx$$

اثبات

$$\sigma^{۲} \geq \int_{-\infty}^{\mu-k\sigma} k^{۲}\sigma^{۲} f(x)dx + \int_{\mu+k\sigma}^{+\infty} k^{۲}\sigma^{۲} f(x)dx$$

با شرط $\sigma^{۲} \neq 0$

$$\frac{۱}{k^{۲}} \geq \int_{-\infty}^{\mu-k\sigma} f(x)dx + \int_{\mu+k\sigma}^{+\infty} f(x)dx$$

پس

$$P(|X-\mu| \geq k\sigma) \leq \frac{۱}{k^{۲}}$$

و یا

$$P(|X-\mu| < k\sigma) \geq ۱ - \frac{۱}{k^{۲}}$$

چبیشف

$$P(|X-\mu| < k\sigma) \geq ۱ - \frac{۱}{k^۲}$$

من باب مثال

- احتمال قرار گرفتن X در فاصلهٔ ۲σ میانگین حداقل $\frac{۳}{۴} = ۱ - \frac{۱}{۲^۲}$
- احتمال قرار گرفتن X در فاصلهٔ ۳σ میانگین حداقل $\frac{۸}{۹} = ۱ - \frac{۱}{۳^۲}$
- احتمال قرار گرفتن X در فاصلهٔ ۵σ میانگین حداقل $\frac{۲۴}{۲۵} = ۱ - \frac{۱}{۵^۲}$

σ کنترل‌کنندهٔ پراکندگی توزیع متغیر تصادفی

نامساوی چبیشف

اگر X متغیر تصادفی برنولی (n,p) باشد. کران بالای $P(X \geq \alpha n)$؟ فرض p<α<1

- $P = \frac{1}{2}, \alpha = \frac{3}{4}$

$$P\{X \geq \alpha n\} = P\{X - np \geq \alpha n - np\} \leq P\{|X - np| \geq \alpha n - np\} \leq \frac{\sigma^۲}{(\alpha n - np)^۲}$$

$$= \frac{p(1-p)}{n(\alpha - p)^۲}$$

$$P\left\{X \geq \frac{3n}{4}\right\} \leq \frac{4}{n}$$

مثال

$$f(x) = \begin{cases} ۶۳۰x^{۴}(۱-x)^{۴}, & 0 < x < ۱ \\ 0 & \text{سایرجاها} \end{cases}$$

احتمال قرار گرفتن X در فاصلهٔ ۲σ از میانگین و مقایسه با کران پائینی حاصل از قضیهٔ چبیشف

$$\mu = \frac{۱}{۲} \text{ میانگین و} \sigma^{۲} = \frac{۱}{۴۴} \Longrightarrow \text{انحراف } \sigma = \sqrt{\frac{۱}{۴۴}} = ۰.۱۵$$

احتمال بودن در فاصلهٔ مدنظر در بازهٔ ۰.۲ و ۰.۸

$$P(.۲ < X < 0.۸) = \int_{۰.۲}^{۰.۸} ۶۳۰x^{۴}(۱-x)^{۴}dx = ۰.۹۶$$

نتیجه حاصل از قضیهٔ چبیشف ۰.۷۵

کران چرنف

متغیر تصادفی X و مقدار حقیقی $a \in \mathrm{R}$

$$P(X \geq a) = P\left(e^{tX} \geq e^{ta}\right), \forall t > 0$$
$$P(X \leq a) = P\left(e^{tX} \geq e^{ta}\right), \forall t < 0$$

طبق نامساوی مارکوفی

$$P(X \geq a) = P\left(e^{tX} \geq e^{ta}\right) \leq \frac{E[e^{tX}]}{e^{sa}}$$

کران چرنف

$$P(X \geq a) \leq e^{-ta}\Phi_{\mathrm{X}}(t), \forall t > 0$$
$$P(X \leq a) \leq e^{-ta}\Phi_{\mathrm{X}}(t), \forall t < 0$$

کران چرنف

اگر X متغیر تصادفی دوجمله‌ای (n,p) باشد. کران بالای $P(X \geq \alpha \mathrm{n})$؟ فرض p<$\alpha$<1

- $P = \frac{1}{2}, \alpha = \frac{3}{4}$

$$\Phi_X(t) = (pe^t + 1 - p)^n$$

$$P\{X \geq \alpha n\} \leq \min_{t>0} e^{-ta} \Phi_X(t) = \min_{t>0} e^{-ta} (pe^t + 1 - p)^n$$

$$P(X \geq \alpha n) \leq \left(\frac{1-p}{1-\alpha}\right)^{(1-\alpha)n} \left(\frac{p}{\alpha}\right)^{\alpha n}.$$

$$P(X \geq \frac{3}{4} n) \leq \left(\frac{16}{27}\right)^{\frac{n}{4}}$$

کران‌های مارکوفی وچبیشف و چرنف

مارکوف ضعیف‌ترین کران

چبیشف کران قوی‌تر و میل به صفر با افزایش **n**

چرنف قوی‌ترین کران و میل به صفر با سرعت نمایی

$$P(X \geq \frac{3n}{4}) \leq \frac{2}{3}$$
$$P(X \geq \frac{3n}{4}) \leq \frac{4}{n}$$
$$P(X \geq \frac{3n}{4}) \leq (\frac{16}{27})^{\frac{n}{4}}$$

نامساوی کوشی شوارتز

معرفی در جبر خطی

معتبر در متغیرهای تصادفی

$$|E[XY]| \leq \sqrt{E[X^{۲}]E[Y^{۲}]}$$

نامساوی ینسن JENSEN

یادآوری

$$Var(X) = E\left[X^{۲}\right] - (E[X])^{۲} \geq 0$$
$$E\left[X^{۲}\right] \geq (E[X])^{۲}$$

تعریف تابع کوژ

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)g(y)$$

نامساوی ینسن هر تابع کوژ

$$E[g(X)] \geq g(E[X])$$

قضایای حدی

مدل‌های همگرائی متغیرهای تصادفی

قضایای حدی

- از نتایج بنیادین نظریه احتمال

نمونه‌های آن

- قانون اعداد بزرگ
  - همگرایی میانگین تعداد زیادی از متغیرهای تصادفی iid به امید ریاضی
- قضیه حد مرکزی
  - جمع تعداد زیادی متغیر تصادفی برابر تقریبی از توزیع نرمال

میانگین نمونه

اگر $X_1, X_2, \ldots, X_n$ توزیع مستقل و یکسان داشته باشند

- ، آن‌گاه متغیر $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ میانگین نمونه

میانگین نمونه

اگر $X_۱,X_۲,\ldots,X_n$ توزیع مستقل و یکسان داشته باشند

- ، آن‌گاه متغیر $\bar{X} = \frac{\sum_{i=۱}^{n} X_i}{n}$ میانگین نمونه

قضیه

اگر $X_۱,X_۲,\ldots,X_n$ توزیع مستقل و یکسان با میانگین μ و وردائی $\sigma^۲$ باشند، آن‌گاه

- $\mathrm{E}[\bar{X}] = \mu$
- $\mathrm{Var}[\bar{X}] = \sigma^۲/n$
- $\mathrm{Cov}(\bar{X},X_i - \bar{X}) = 0$

اثبات

میانگین نمونه

- اثبات

$$\mathrm{E}[\bar{X}]$$

میانگین نمونه

اثبات-

$$\mathrm{E}[\bar{X}] = \frac{۱}{n}\sum_{i=۱}^{n} E[X_i]$$

میانگین نمونه

-اثبات

$$\mathrm{E}[\bar{X}] = \frac{۱}{n}\sum_{i=۱}^{n} E[X_i] = \mu$$

میانگین نمونه

اثبات-

$$\mathrm{E}[\bar{X}] = \frac{۱}{n}\sum_{i=۱}^{n} E[X_i] = \mu$$

$$\mathrm{Var}[\bar{X}] = \left(\frac{۱}{n}\right)^{۲} Var\left(\sum_{i=۱}^{n} X_i\right) = \left(\frac{۱}{n}\right)^{۲} \sum_{i=۱}^{n} Var(X_i) = \frac{\sigma^{۲}}{n}$$

میانگین نمونه

اثبات-

$$\mathrm{E}[\bar{X}] = \frac{۱}{n}\sum_{i=۱}^{n} E[X_i] = \mu$$

$$\mathrm{Var}[\bar{X}] = \left(\frac{۱}{n}\right)^{۲} Var\left(\sum_{i=۱}^{n} X_i\right) = \left(\frac{۱}{n}\right)^{۲} \sum_{i=۱}^{n} Var(X_i) = \frac{\sigma^{۲}}{n}$$

$$\mathrm{Cov}(\bar{X}, X_i - \bar{X}) = \mathrm{Cov}(\bar{X}, X_i) - \mathrm{Cov}(\bar{X}, \bar{X}) = \frac{۱}{\mathrm{n}}\mathrm{Cov}\left(X_i + \sum_{j \neq i}^{n} X_j\,, X_i\right) - \mathrm{Var}[\bar{X}]$$

$$= \frac{۱}{\mathrm{n}}\mathrm{Cov}(X_i, X_i) + \frac{۱}{\mathrm{n}}\mathrm{Cov}\left(\sum_{j \neq i}^{n} X_j\,, X_i\right) - \mathrm{Var}[\bar{X}] = \frac{\sigma^{۲}}{n} - \frac{\sigma^{۲}}{n} = 0$$

توزیع نرمال استانده

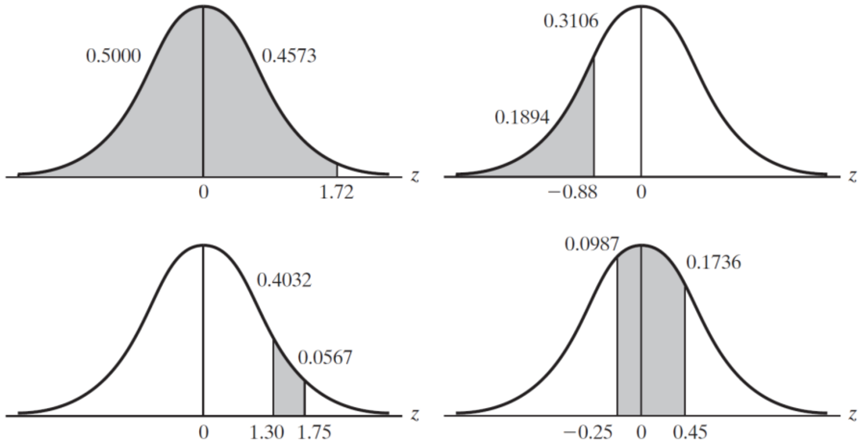
میانگین صفر و انحراف معیار ۱

نمایش با Z

Table III: Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4988
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Also, for $z = 4.0$, 5.0, and 6.0, the probabilities are 0.49997, 0.4999997, and 0.499999999.



مثال- متغیر تصادفی با توزیع نرمال استانده

احتمال مقدار کمتر از ۱.۷۲
- یافتن مقدار متناظر در جدول و افزودن ۰.۵۰۰۰ به آن
- ۰.۵۰۰۰ + ۰.۴۵۷۳ = ۰.۹۵۷۳

احتمال مقدار کمتر از −۰.۸۸
- یافتن مقدار متناظر در جدول و تفاضل آن از ۰.۵۰۰۰
- ۰.۵۰۰۰ − ۰.۳۱۰۶ = ۰.۱۸۹۴

احتمال مقدار بین ۱.۳۰ و ۱.۷۵
- تفاضل بزرگتر از کوچکتر
- ۰.۴۵۹۹ − ۰.۴۰۳۲ = ۰.۰۵۶۷

احتمال مقدار بین −۰.۲۵ و ۰.۴۵
- جمع
- ۰.۰۹۸۷ + ۰.۱۷۳۶ = ۰.۲۷۲۳

Table III: Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4988
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Also, for $z = 4.0$, 5.0, and 6.0, the probabilities are 0.49997, 0.4999997, and 0.499999999.

قانون قوی اعداد بزرگ

اگر توزیع ... $X_۱, X_۲,$ مستقل و یکسان داشته باشند و $\mathrm{E}[X_i] = \mu$ با احتمال ۱، آن‌گاه متغیر $\bar{X} = \frac{\sum_{i=۱}^{n} X_i}{n}$ میانگین نمونه خوانده می‌شود.

$$\frac{X_۱ + X_۲ + \cdots X_n}{n} \to \mu \text{ با } n \to \infty$$

قضیهٔ حد مرکزی (یا قضیهٔ مرکزی حد)

قضیه

اگر $X_۱, X_۲, \ldots, X_n$ دنبالهٔ مستقلی از متغیرهای تصادفی مستقل و یکسان با میانگین μ و وردائی $\sigma^۲$ باشند، آن‌گاه

$$\frac{X_۱ + X_۲ + \cdots X_n - n\mu}{\sigma\sqrt{n}}$$

با $n \to \infty$ به توزیع نرمال استانده میل می‌کند. به دیگر سخن

$$P\left\{\frac{X_۱ + X_۲ + \cdots X_n - n\mu}{\sigma\sqrt{n}} \le a\right\} \to \frac{۱}{\sqrt{۲\pi}} \int_{-\infty}^{a} e^{-\frac{x^۲}{۲}} dx \text{ با } n \to \infty$$

یا با نمایش دیگر $\bar{X} = \frac{\sum_{i=۱}^{n} X_i}{n}$ و $n \to \infty$، $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ توزیع نرمال استاندارد است.

قضیه دوموآور-۱۷۱۸

Xمتغیر تصادفی دوجمله‌ای با پارامترهای n و $\frac{۱}{۲}$ باشد، آن‌گاه به ازای هردو عدد a و b و $a<b$ داریم

$$\lim_{n\to\infty} P\left(a < \frac{X-\frac{۱}{۲}n}{\frac{۱}{۲}\sqrt{n}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$$

$$E[X] = \frac{۱}{۲}n$$

$$\sigma_X = \frac{۱}{۲}\sqrt{n}$$

قضیه دوموآور-لاپلاس-۱۸۱۲

X متغیر تصادفی دوجمله‌ای با پارامترهای n و p باشدبرای هردو عدد a و b و $a<b$ داریم

$$\lim_{n\to\infty} P\left(a < \frac{X-np}{\sqrt{np(1-p)}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$$

$$E[X] = np$$

$$\sigma_X = \sqrt{np(1-p)}$$

تقریبی مناسب برای هر n و p به شرط $np(1-p) \geq$ ۱۰

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

احتمال حداقل میانگین ۲۰۴ برای نمونه تصادفی به اندازه ۳۶ مورد

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

احتمال حداقل میانگین ۲۰۴ برای نمونه تصادفی به اندازه ۳۶ مورد

بنابر چند قضیه قبل

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

احتمال حداقل میانگین ۲۰۴ برای نمونه تصادفی به اندازه ۳۶ مورد

بنابر چند قضیه قبل

- میانگین توزیع $\bar{X}$ برابر ۲۰۰ $= \mu_{\bar{X}}$

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

احتمال حداقل میانگین ۲۰۴ برای نمونه تصادفی به اندازه ۳۶ مورد

بنابر چند قضیه قبل

- میانگین توزیع $\bar{X}$ برابر ۲۰۰ $= \mu_{\bar{X}}$
- انحراف معیار آن برابر با ۲.۵ $= \sigma_{\bar{X}} = \frac{۱۵}{\sqrt{۳۶}}$

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

احتمال حداقل میانگین ۲۰۴ برای نمونه تصادفی به اندازه ۳۶ مورد

بنابر چند قضیه قبل

- میانگین توزیع $\bar{X}$ برابر ۲۰۰ $= \mu_{\bar{X}}$
- انحراف معیار آن برابر با ۲.۵ $= \sigma_{\bar{X}} = \frac{۱۵}{\sqrt{۳۶}}$
- بر اساس قضیه مرکزی حد توزیع تقریبا نرمال ۱.۶ $= Z = \frac{۲۰۴ - ۲۰۰}{۲.۵}$

مثال

متغیر تصادفی با میانگین ۲۰۰ و انحراف معیار ۱۵

احتمال حداقل میانگین ۲۰۴ برای نمونه تصادفی به اندازه ۳۶ مورد

بنابر چند قضیه قبل

- میانگین توزیع $\bar{X}$ برابر $\mu_{\bar{X}} = ۲۰۰$
- انحراف معیار آن برابر با $\sigma_{\bar{X}} = \frac{۱۵}{\sqrt{۳۶}} = ۲.۵$
- بر اساس قضیه مرکزی حد توزیع تقریبا نرمال $Z = \frac{۲۰۴-۲۰۰}{۲.۵} = ۱.۶$

$$P(\bar{X} \geq ۲۰۴) = P(Z \geq ۱.۶) = ۰.۵۰۰۰ - ۰.۴۴۵۲ = ۰.۰۵۴۸$$

محاسبهٔ امید با شرطی بودن

$$E[X] = E\big[E[X|Y]\big]$$

گسسته

$$E[X] = \sum_{y} \big[E[X|Y = y]\big] P\{Y = y\}$$

پیوسته

$$E[X] = \int_{-\infty}^{+\infty} \big[E[X|Y = y]\big] f_Y(y)$$

محاسبهٔ امید با شرطی بودن- اثبات حالت گسسته

$$E\big[E[X|Y]\big] = \sum_y \big[E[X|Y = y]\big] P\{Y = y\}$$
$$= \sum_y \sum_x xP\{X = x|Y = y\}\, P\{Y = y\}$$
$$= \sum_y \sum_x xP\{X = x, Y = y\}$$
$$= \sum_x x \sum_y P\{X = x, Y = y\}$$
$$= \sum_x xP\{X = x\}$$
$$= E[X]$$

محاسبهٔ وردائی با شرطی بودن

$$Var[X] = E\big[Var[X|Y]\big] + Var[E[X|Y]]$$

استنتاج آماری

در زندگی واقعی

- کار با داده‌های تحت تأثیر تصادف
- نیاز به استخراج اطلاع جهت نتیجه
- نتیجه‌گیری از داده

استنتاج آماری

در زندگی واقعی

- کار با داده‌های تحت تأثیر تصادف
- نیاز به استخراج اطلاع جهت نتیجه
- نتیجه‌گیری از داده

منبع تصادف

استنتاج آماری

در زندگی واقعی

- کار با داده‌های تحت تأثیر تصادف
- نیاز به استخراج اطلاع جهت نتیجه
- نتیجه‌گیری از داده

منبع تصادف

- عدم دسترسی به تمامی جمعیت
  - انتخاب مجموعه‌ای تصادفی از جمعیت و پرسش از آنها
  - تصادف ناشی از نمونه‌برداری

استنتاج آماری

در زندگی واقعی

- کار با داده‌های تحت تأثیر تصادف
- نیاز به استخراج اطلاع جهت نتیجه
- نتیجه‌گیری از داده

منبع تصادف

- عدم دسترسی به تمامی جمعیت
  - انتخاب مجموعه‌ای تصادفی از جمعیت و پرسش از آنها
  - تصادف ناشی از نمونه‌برداری
- ناشی از نوفه

استنتاج آماری

در زندگی واقعی

- کار با داده‌های تحت تأثیر تصادف
- نیاز به استخراج اطلاع جهت نتیجه
- نتیجه‌گیری از داده

منبع تصادف

- عدم دسترسی به تمامی جمعیت
  - انتخاب مجموعه‌ای تصادفی از جمعیت و پرسش از آنها
  - تصادف ناشی از نمونه‌برداری
- ناشی از نوفه

منجر به استنتاج آماری

- مجموعه روش‌هائی که از داده‌های مخلوط با تصادف سعی بر استخراج نتیجه‌گیری می‌کنند

استنتاج آماری

نیاز به استفاده از دانش احتمالی

کار با داده حقیقی

- عدم روشن بودن توزیع داده واقعی

روش‌های اصلی

- بسامدی
- بیزی

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

فرض بر ثابت بودن (تصادفی نبودن) مقدار θ

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

فرض بر ثابت بودن (تصادفی نبودن) مقدار θ

الگوریتم

- جمع‌آوری چند نمونه جهت تخمین θ
  - دریافت نمونه تصادفی X1، X2، ...، Xn

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

فرض بر ثابت بودن (تصادفی نبودن) مقدار θ

الگوریتم

- جمع‌آوری چند نمونه جهت تخمین θ
  - دریافت نمونه تصادفی X1، X2، ...، Xn
- تعریف تخمین‌گر نقطه‌ای $\hat{\theta}$ جهت تخمین θ به طوری که $\hat{\theta} = h(X_1, \cdots, X_n)$

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

فرض بر ثابت بودن (تصادفی نبودن) مقدار θ

الگوریتم

- جمع‌آوری چند نمونه جهت تخمین θ
  - دریافت نمونه تصادفی X1، X2، ...، Xn
- تعریف تخمین‌گر نقطه‌ای $\hat{\theta}$ جهت تخمین θ به طوری که $\hat{\theta} = h(X_1, \cdots, X_n)$
  - مثال $\hat{\theta} = \bar{X} = \frac{X_1 + \cdots + X_n}{n}$

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

فرض بر ثابت بودن (تصادفی نبودن) مقدار θ

الگوریتم

- جمع‌آوری چند نمونه جهت تخمین θ
  - دریافت نمونه تصادفی X1، X2، ...، Xn
- تعریف تخمین‌گر نقطه‌ای $\hat{\theta}$ جهت تخمین θ به طوری که $\hat{\theta} = h(X_1, \cdots, X_n)$
  - مثال $\hat{\theta} = \bar{X} = \frac{X_1 + \cdots + X_n}{n}$
- وجود بی‌نهایت تخمین‌گر ممکن برای θ

تخمین نقطه‌ای

تلاش بر تخمین پارامتر مجهول θ

- مثال $\theta = E[X]$

فرض بر ثابت بودن (تصادفی نبودن) مقدار θ

الگوریتم

- جمع‌آوری چند نمونه جهت تخمین θ
  - دریافت نمونه تصادفی X1، X2، ...، Xn
- تعریف تخمین‌گر نقطه‌ای $\hat{\theta}$ جهت تخمین θ به طوری که $\hat{\theta} = h(X_۱, \cdots, X_n)$
  - مثال $\hat{\theta} = \bar{X} = \frac{X_۱ + \cdots + X_n}{n}$
- وجود بی‌نهایت تخمین‌گر ممکن برای θ
  - چگونگی اطمینان از مناسب بودن تخمین‌گر
  - چگونگی مقایسهٔ تخمین‌گرهای متفاوت

ارزیابی تخمین‌گرها

سه خاصیت مطلوب برآورده‌سازهای نقطه‌ای $\hat{\theta} = h(X_1, \cdots, X_n)$ از θ

- اریبی $\mathrm{B}(\hat{\theta}) = E[\hat{\theta}] - \theta$
- میانگین مربع خطا
- سازگاری

ارزیابی تخمین‌گرها

اریبی $\mathrm{B}(\hat{\theta}) = E[\hat{\theta}] - \theta$

- در پی برابری اریب با مقدار صفر یا $E[\hat{\theta}] = \theta$
- به معنای به طور میانگین نزدیکی $\hat{\theta}$ به θ

مثال فرض $\hat{\theta} = \bar{X} = \frac{X_1 + \cdots + X_n}{n}$ اریب یا نااریب؟

$$\mathrm{B}(\hat{\theta}) = E[\hat{\theta}] - \theta = E[\bar{X}] - \theta = E[X_i] - \theta = 0$$

- پس نااریب
- اما لزوما مناسب نیست.

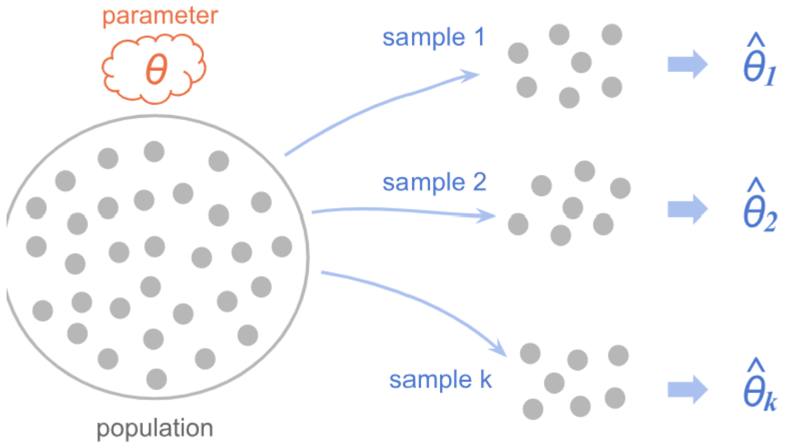
ارزیابی تخمین‌گرها

میانگین مربع خطا MSE تخمین‌گر $\hat{\theta}$ برابر $MSE[\hat{\theta}] = E\left[(\hat{\theta} - \theta)^{۲}\right]$

- $\theta - \hat{\theta}$ برابر خطای تخمین θ با $\hat{\theta}$
- ممخ معیاری از میزان فاصلهٔ بین $\hat{\theta}$ و θ
- مقدار کمتر معادل تخمین مناسب‌تر

مثال

ارزیابی تخمین‌گرها- میانگین مربع خطا



ارزیابی تخمین‌گرها- میانگین مربع خطا

$$
\begin{aligned}
(\hat\theta-\theta)^2 &= \left(\hat\theta-\mathbb{E}(\hat\theta)+\mathbb{E}(\hat\theta)-\theta\right)^2\\
&= (\hat\theta-\mu_{\hat\theta}+\mu_{\hat\theta}-\theta)^2\\
&= (\underbrace{\hat\theta-\mu_{\hat\theta}}_{a}+\underbrace{\mu_{\hat\theta}-\theta}_{b})^2\\
&= a^2+b^2+2ab\\
\implies \mathbb{E}\left[(\hat\theta-\theta)^2\right] &= \mathbb{E}[a^2+b^2+2ab]
\end{aligned}
$$

ارزیابی تخمین‌گرها- میانگین مربع خطا

امکان تجزیه MSE به

$$
\begin{aligned}
\mathbb{E}\left[(\hat{\theta}-\theta)^2\right] &= \mathbb{E}[a^2+b^2+2ab] \\
&= \mathbb{E}(a^2)+\mathbb{E}(b^2)+2\mathbb{E}(ab) \\
&= \mathbb{E}[(\hat{\theta}-\mu_{\hat{\theta}})^2]+\mathbb{E}[(\mu_{\hat{\theta}}-\theta)^2]+2\mathbb{E}(ab)
\end{aligned}
$$

ارزیابی تخمین‌گرها- میانگین مربع خطا

اما

$$\mathbb{E}(ab) = \mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}})(\mu_{\hat{\theta}} - \theta)] = 0$$

ارزیابی تخمین‌گرها- میانگین مربع خطا

در نتیجه

$$\mathrm{MSE}(\hat{\theta}) = \mathbb{E}\left[(\hat{\theta} - \theta)^2\right]$$

$$= \mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}})^2] + \mathbb{E}[(\mu_{\hat{\theta}} - \theta)^2]$$

$$= \mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}})^2] + \mathbb{E}[(\mu_{\hat{\theta}} - \theta)]^2$$

$$= \underbrace{\mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}})^2]}_{\text{Variance}} + (\underbrace{\mu_{\hat{\theta}} - \theta}_{\text{Bias}})^2$$

$$= \mathrm{Var}(\hat{\theta}) + \mathrm{Bias}^2(\hat{\theta})$$

ارزیابی تخمین‌گرها- میانگین مربع خطا

$$MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^{۲}\right] = Var[\hat{\theta}] + \mathrm{B}(\hat{\theta})^{۲}$$

اریب: تمایل $\hat{\theta}$ به فرا- یا فرو- تخمین θ

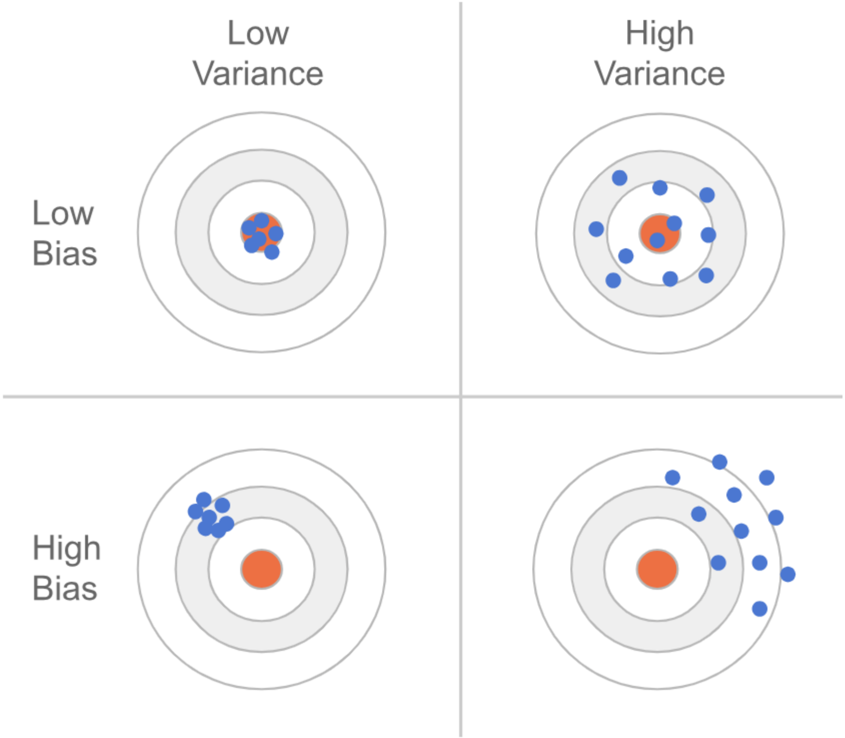
وردائی: میانگین پراکندگی تخمین‌گرها حول مقدار میانگین خود $E[\hat{\theta}]$

سخن کوتاه

- میانگین مربع خطا تخمین‌گر برابر جمع اندازه‌گیری میزان دوری تخمین‌گر به طور میانگین از مقدار امید ریاضی‌اش و اندازه‌گیری پراکندگی تخمین‌گر

ارزیابی تخمین‌گرها- میانگین مربع خطا

نمونه موارد اریبی و وردائی



ارزیابی تخمین‌گرها- میانگین مربع خطا

مثال- فرض X1، X2، ...، Xn نمونه تصادفی از توریع با میانگین $E[X_i] = \theta$ و وردائی $Var(X_i) = \sigma^۲$، با دو تخمین

- ۱- $\widehat{\theta_۱} = X_۱$
- ۲- $\widehat{\theta_۲} = \bar{X} = \frac{X_۱+\cdots+X_n}{n}$

حل-

$$MSE(\widehat{\theta_۱}) = E\left[\left(\widehat{\theta_۱} - \theta\right)^۲\right] = E[(X_۱ - E[X_۱])^۲] = Var[X_۱] = \sigma^۲$$

$$MSE(\widehat{\theta_۲}) = E\left[\left(\widehat{\theta_۲} - \theta\right)^۲\right] = E[(\bar{X} - \theta)^۲] = Var[(\bar{X} - \theta)] + E[(\bar{X} - \theta)]^۲ = Var[\bar{X}] = \frac{\sigma^۲}{n}$$

$$\forall n > 1:\ MSE(\widehat{\theta_۱}) > MSE(\widehat{\theta_۲})$$

ارزیابی تخمین‌گرها- میانگین مربع خطا

سخن کوتاه،

$$MSE(\hat{\theta}) = E\left[\left(\hat{\theta} - \theta\right)^۲\right]$$

$$= Var[\hat{\theta} - \theta] + E[\hat{\theta} - \theta]^۲$$

$$= Var[\hat{\theta}] + B(\hat{\theta})^۲$$

به طوری که $B(\hat{\theta}) = E[\hat{\theta}] - \theta$

ارزیابی تخمین‌گرها

سازگاری

تعریف- فرض $\widehat{\theta_1}$، $\widehat{\theta_2}$، ...، $\widehat{\theta_n}$، ... دنباله‌ای از تخمین‌سازهای نقطه‌ای از θ باشند. آن‌گاه $\widehat{\theta_n}$ تخمین‌ساز «سازگاری» از θ است، اگر

$$\lim_{n\to\infty} P\left(\left|\widehat{\theta_n}-\theta\right| \geq \epsilon\right) = 0 \,, \forall \epsilon > 0$$

ارزیابی تخمین‌گرها

سازگاری

تعریف- فرض $\widehat{\theta_۱}$، $\widehat{\theta_۲}$، ...، $\widehat{\theta_n}$، ... دنباله‌ای از تخمین‌سازهای نقطه‌ای از θ باشند. آن‌گاه $\widehat{\theta_n}$ تخمین‌ساز «سازگاری» از θ است، اگر

$$\lim_{n\to\infty} P\left(\left|\widehat{\theta_n}-\theta\right| \geq \epsilon\right) = 0\,, \forall \epsilon > 0$$

قضیه- فرض $\widehat{\theta_۱}$، $\widehat{\theta_۲}$، ...، $\widehat{\theta_n}$، ... دنباله‌ای از تخمین‌سازهای نقطه‌ای از θ باشند. اگر

$$\lim_{n\to\infty} MSE\left(\widehat{\theta_n}\right) = 0$$

آن‌گاه $\widehat{\theta_n}$ تخمین‌ساز «سازگاری» از θ است.

تخمین‌ساز نقطه‌ای میانگین و وردائی

تعریف- فرض X1، X2، ...، Xn نمونه تصادفی از توزیع

- با میانگین $E[X_i] = \mu < \infty$ و
- با وردائی $0 < Var(X_i) = \sigma^{۲} < \infty$
- «وردائی نمونه» نمونه تصادفی مذکور برابر است با

$$S^{۲} = \frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i - \bar{X})^{۲} = \frac{۱}{n-۱}(\sum_{i=۱}^{n} {X_i}^{۲} - n\bar{X})$$

به طوری که وردائی نمونه تخمین‌ساز نااریب $\sigma^{۲}$ است. همچنین انحراف معیار نمونه برابر $S = \sqrt{S^{۲}}$.

S تخمین‌سازی اریب از σ است.

تخمین بیشینه درست‌نمائی MAXIMUM LIKELIHOOD ESTIMATION

راه‌حلی روشمند جهت تخمین پارامترها

مثال- کوزه‌ای حاوی سه توپ و هر یک یا سرخ یا کبود ولی عدم اطلاع بیشتر

فرض تعداد توپ‌هایی کبود برابر θ

- مقادیر ممکن آن ۰، ۱، ۲، یا ۳.

اجازهٔ انتخاب چهارتوپ از کوزه با جاگذاری.

تعریف مت‌های X1، X2، X3، X4

$$X_i = \begin{cases} ۱, & \text{توپ } i \text{ کبود} \\ 0, & \text{توپ } i \text{ سرخ} \end{cases}$$

تخمین بیشینه درست‌نمائی

$$X_i = \begin{cases} ۱, & \text{توپ } i \text{ کبود} \\ 0, & \text{توپ } i \text{ سرخ} \end{cases}$$

X_i دارای توزیع مستقل و یکسان و $X_i \sim Bernoulli\left(\frac{\theta}{3}\right)$

پس از انجام آزمایش، مشاهدهٔ مقادیر X_i به ترتیب

$x_۱ = ۱, x_۲ = 0, x_۳ = ۱, x_۴ = ۱$ یا $(x_۱, x_۲, x_۳, x_۴) = (۱, 0, ۱, ۱)$

به ازای چه احتمالی از θ مقدار احتمال مشاهدهٔ فوق بیشینه است؟

تخمین بیشینه درست‌نمائی

$x_۱ = ۱, x_۲ = 0, x_۳ = ۱, x_۴ = ۱$ یا $(x_۱, x_۲, x_۳, x_۴) = (۱,0,۱,۱)$

به ازای چه احتمالی از θ مقدار احتمال مشاهدهٔ فوق بیشینه است؟

- حل- چون $X_i \sim Bernoulli\left(\frac{\theta}{3}\right)$ پس

$$P_{X_i}(x) = \begin{cases} \frac{\theta}{۳}, x = ۱ \\ ۱ - \frac{\theta}{۳}, x = 0 \end{cases}$$

- داریم $P_{X_۱ X_۲ X_۳ X_۴}(x_۱, x_۲, x_۳, x_۴) = \frac{\theta}{۳}\left(۱ - \frac{\theta}{۳}\right)\frac{\theta}{۳}\frac{\theta}{۳} = \left(\frac{\theta}{۳}\right)^۳\left(۱ - \frac{\theta}{۳}\right)$

تخمین بیشینه درست‌نمائی

داریم $P_{X_۱X_۲X_۳X_۴}(x_۱,x_۲,x_۳,x_۴) = \frac{\theta}{۳}\left(۱-\frac{\theta}{۳}\right)\frac{\theta}{۳}\frac{\theta}{۳} = \left(\frac{\theta}{۳}\right)^۳\left(۱-\frac{\theta}{۳}\right)$

جدول مقادیر مختلف

θ	$P_{X_1X_2X_3X_4}(1,0,1,1;\theta)$
0	0
1	0.0247
2	0.0988
3	0

پاسخ $\theta = ۲$

تخمین درست‌نمائی

فرض تعریف- فرض X1، X2، ...، Xn نمونه تصادفی از توزیع با پارامتر θ باشد. در صورت مشاهدهٔ مقادیر X1=x1، X2=x2، ...، Xn=xn

اگر X_iها گسسته باشند، آن‌گاه تابع درست‌نمائی برابر است با

$$L(x_۱, x_۲, \cdots, x_n; \theta) = P_{X_۱, X_۲, \cdots, X_n}(x_۱, x_۲, \cdots, x_n; \theta)$$

اگر X_iها پیوسته باشند، آن‌گاه تابع درست‌نمائی برابر است با

$$L(x_۱, x_۲, \cdots, x_n; \theta) = f_{X_۱, X_۲, \cdots, X_n}(x_۱, x_۲, \cdots, x_n; \theta).$$

در بعضی از مسائل حل لگاریتم درست‌نمائی راحتتر است و به صورت زیر نمایش داده می‌شود

$$\ln L(x_۱, x_۲, \cdots, x_n; \theta)$$

تخمین بیشینه درست‌نمائی

فرض تعریف- فرض X1، X2، ...، Xn نمونه تصادفی از توزیع با پارامتر θ باشد. در صورت مشاهدهٔ مقادیر X1=x1، X2=x2، ...، Xn=xn آن‌گاه تخمین بیشینهٔ درست‌نمائی θ را با $\hat{\theta}_{ML}$ نمایش می‌دهیم برابر با مقداری از θ است که تابع درستنمایی $L(x_1, x_2, \cdots, x_n; \theta)$ را بیشینه می‌کند. تخمین‌گر بیشینه درست‌نمائی از پارامتر θ متغیر تصادفی است که $\hat{\theta}_{ML}$ برمی‌گرداند.

تخمین بیشینه درست‌نمائی

مثال « $(۳,\theta)$دوجمله‌ای$X_i\sim$» و مشاهدهٔ $(۱,۳,۲,۲) = (x_۱,x_۲,x_۳,x_۴)$.

تابع دوجمله‌ای برابر است با

$$P_{X_i}(x;\theta) = \binom{3}{x}\theta^x(1-\theta)^{3-x}$$

پس

$$\begin{aligned} L(x_1,x_2,x_3,x_4;\theta) &= P_{X_1X_2X_3X_4}(x_1,x_2,x_3,x_4;\theta) \\ &= P_{X_1}(x_1;\theta)P_{X_2}(x_2;\theta)P_{X_3}(x_3;\theta)P_{X_4}(x_4;\theta) \\ &= \binom{3}{x_1}\binom{3}{x_2}\binom{3}{x_3}\binom{3}{x_4}\theta^{x_1+x_2+x_3+x_4}(1-\theta)^{12-(x_1+x_2+x_3+x_4)} \end{aligned}$$

با مقادیر مشاهده شده داریم

$$\begin{aligned} L(1,3,2,2;\theta) &= \binom{3}{1}\binom{3}{3}\binom{3}{2}\binom{3}{2}\theta^8(1-\theta)^4 \\ &= 27 \qquad \theta^8(1-\theta)^4. \end{aligned}$$

تخمین بیشینه درست‌نمائی

مثال « $(۳,\theta)$دوجمله‌ای$\sim X_i$» و مشاهدۀ $(۱,۳,۲,۲) = (x_۱,x_۲,x_۳,x_۴)$.

تابع درست‌نمایی برابر

$$L(1,3,2,2;\theta) = 27 \qquad \theta^8(1-\theta)^4$$

مشتق‌گیری جهت یافتن مقدار بیشینه‌سازی تابع مذکور

$$\frac{dL(1,3,2,2;\theta)}{d\theta} = 27[\qquad 8\theta^7(1-\theta)^4 - 4\theta^8(1-\theta)^3]$$

در نتیجه $\hat{\theta}_{\text{بد}} = \frac{۲}{۳}$

تخمین بیشینه درست‌نمائی

مثال « $X_i \sim$ نمائی(θ)» و مشاهدهٔ $(x_۱,x_۲,x_۳,x_۴) = (۱.۲۳,۳.۳۲,۱.۹۸,۲.۱۲)$.

تابع نمایی برابر است با

$$f_{X_i}(x;\theta) = \theta e^{-\theta x} u(x)$$

پس

$$\begin{aligned} L(x_1,x_2,x_3,x_4;\theta) &= f_{X_1X_2X_3X_4}(x_1,x_2,x_3,x_4;\theta) \\ &= f_{X_1}(x_1;\theta)f_{X_2}(x_2;\theta)f_{X_3}(x_3;\theta)f_{X_4}(x_4;\theta) \\ &= \theta^4 e^{-(x_1+x_2+x_3+x_4)\theta}. \end{aligned}$$

با مقادیر مشاهده شده داریم

$$L(1.23,3.32,1.98,2.12;\theta) = \theta^4 e^{-8.65\theta}.$$

تخمین بیشینه درست‌نمائی

مثال « $X_i \sim$ نمائی(θ)» و مشاهدهٔ $(x_۱, x_۲, x_۳, x_۴) = (۱.۲۳, ۳.۳۲, ۱.۹۸, ۲.۱۲)$.

تابع درست‌نمایی برابر

$$L(1.23, 3.32, 1.98, 2.12; \theta) = \theta^4 e^{-8.65\theta}$$

راحتتر بودن کار با لگاریتم

$$\ln L(1.23, 3.32, 1.98, 2.12; \theta) = 4 \ln \theta - 8.65\theta$$

مشتق‌گیری جهت یافتن مقدار بیشینه‌سازی تابع مذکور $\frac{۴}{\theta} - ۸.۶۵ = 0$

در نتیجه $\hat{\theta}_{\text{بد}} = ۰.۴۶$

تخمین بیشینه درست‌نمائی

تمرین تخمین‌گر بیشینه درست‌نمائی MLE موارد زیر را بیابید

- « $X_i \sim$ دوجمله‌ای $(۳,\theta)$ » و مشاهدهٔ $(x_۱, x_۲, \cdots, x_n)$.
- « $X_i \sim$ نمائی (θ) » و مشاهدهٔ $(x_۱, x_۲, \cdots, x_n)$.

تخمین بیشینه درست‌نمائی

مقدار حاصل از تخمین بیشینهٔ درست‌نمائی تابعی از دادهٔ مشاهده شده

- مقدار محاسبه شده با آن متغیری تصادفی است.
- $\hat{\Theta}_{\text{بد}} = \hat{\theta}_{\text{بد}}$ حاصل تخمین درست‌نمائی بیشینه برای داده مشاهده شده

عدم امکان استفاده از مشتق‌گیری جهت بدست آوردن مقدار بیشینه‌ساز

- مقادیر گسسته
- مقادیر حقیقی ولی تابع مقید.

خواص مجانبی تخمین بیشینه درست‌نمائی

$\hat{\Theta}_{\text{بد}}$ تبد MLE پارامتر θ باشد. آن‌گاه

- $\hat{\Theta}_{\text{بد}}$ مجانبی سازگار است، یا $\underset{n\to\infty}{\text{حد}} P\left(\left|\hat{\Theta}_{\text{بد}} - \theta\right| \geq \epsilon\right) = 0, \forall \epsilon > 0$.

- $\hat{\Theta}_{\text{بد}}$ مجانبی نااریب است، یا $\underset{n\to\infty}{\text{حد}} E\left[\hat{\Theta}_{\text{بد}}\right] = \theta$.

- با بزرگتر شدن **n**، $\hat{\Theta}_{\text{بد}}$ به مت نرمال میل می‌کند. یا مت $\frac{\hat{\Theta}_{\text{بد}}-\theta}{\sqrt{Var\left(\hat{\Theta}_{\text{بد}}\right)}}$ به توزیع N(0,1) میل می‌کند.

رگرسیون خطی

در پی یافتن مدل سادهٔ توضیح دهندهٔ روابط بین دو یا چند مت
$(x_i, y_i), i = 1,2,\cdots,n$

معمولا به دنبال مدل‌های ساده

مدل خطی محتملا ساده‌ترین مدل

$$y_i \approx \beta_0 + \beta_1 x_i$$

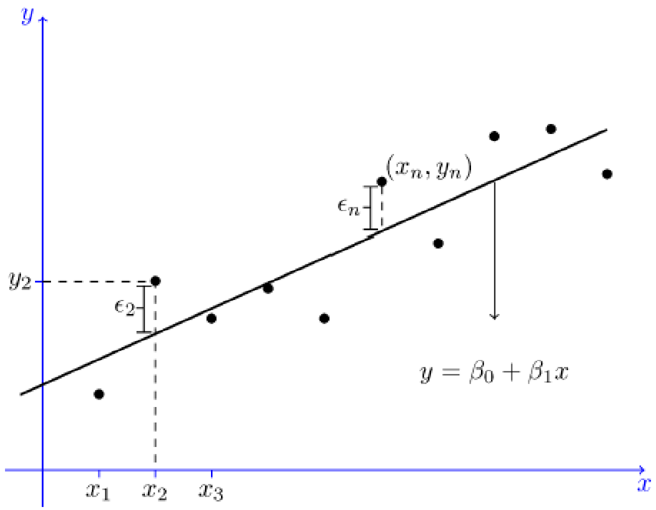
امکان تأثیر عوامل دیگر، پس

$$y_i \approx \beta_0 + \beta_1 x_i + \epsilon_i$$

- ϵ_i به مثابهٔ مت؛ نشان‌دهنده میزان خطا در تقریب
- هدف تخمین β_0 و β_1 با کمترین خطا

خط رگرسیون $\hat{y} = \beta_0 + \beta_1 x$

- بهترین برازش داده‌ها

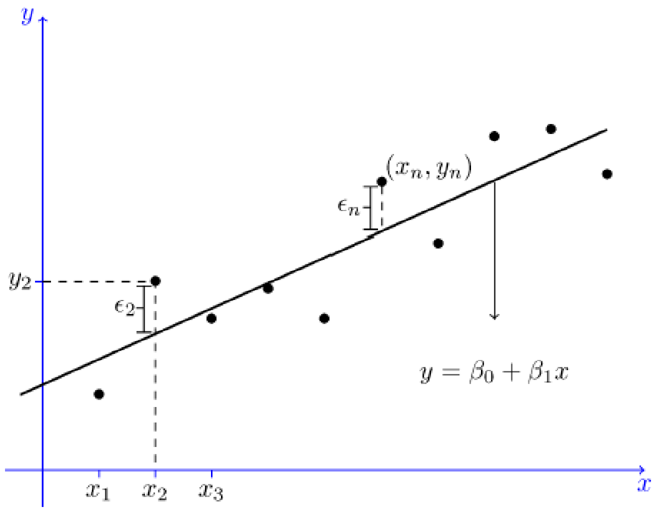


رگرسیون خطی

خط رگرسیون $\hat{y} = \beta_0 + \beta_۱ x$

$Y = \beta_0 + \beta_۱ x + \epsilon$

- به دلیل مت اپسیلون، مت بودن Y
- x معروف به پیش‌بین‌ساز یا متغیر توصیف‌گر
- Y معروف به متغیر پاسخ



مدل سادهٔ رگرسیون خطی

$Y_i = \beta_0 + \beta_۱ x_i + \epsilon_i$

$\epsilon_i \sim N(0,۱)$

فرض بر ثابت ولی مجهول بودن β_0 و $\beta_۱$ و ϵ_i

فرض بر داشتن n مقدار (x_i, y_i)

هدف یافتن بهترین β_0 و $\beta_۱$ مرتبط کنندهٔ x_i با y_i

چند روش حل

روش اول حل مدل ساده رگرسیون خطی

فرض بر مشاهدهٔ x_i-ها از مت X

- پس

$$Y = \beta_0 + \beta_1 X + \epsilon$$

- و$\epsilon \sim N(0,1)$

اعمال امید ریاضی به دو طرف معادله

$$E[Y] = \beta_0 + \beta_1 E[X] + E[\epsilon]$$
$$= \beta_0 + \beta_1 E[X]$$

- در نتیجه

$$\beta_0 = E[Y] - \beta_1 E[X]$$

روش اول حل مدل ساده رگرسیون خطی

محاسبهٔ Cov(X,Y)

$$Cov(X,Y) = \mathrm{Cov}(\mathrm{X}, \beta_0 + \beta_۱ X + \epsilon)$$
$$= \beta_0 \mathrm{Cov}(\mathrm{X}, ۱) + \beta_۱ Cov(X,X) + Cov(X,\epsilon)$$
$$= 0 + \beta_۱ Cov(X,X) + 0$$
$$= \beta_۱ Var(X)$$

- $Cov(X,\epsilon) = 0$ به دلیل استقلال Xو ϵ از یکدیگر

در نتیجه

$$\beta_۱ = \frac{Cov(X,Y)}{Var(X)}, \beta_0 = E[Y] - \beta_۱ E[X]$$

روش اول حل مدل ساده رگرسیون خطی

$$\beta_1 = \frac{Cov(X,Y)}{Var(X)}, \beta_0 = E[Y] - \beta_1 E[X]$$

امکان محاسبهٔ β_1وβ_0 در صورت داشتن E[X]، E[Y]، و $\frac{Cov(X,Y)}{Var(X)}$

با داشتن n مقدار مشاهده شده امکان تخمین مقادیر مزبور

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n},$$

$$\bar{y} = \frac{y_1 + y_2 + \ldots + y_n}{n},$$

$$s_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

$$s_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

روش اول حل مدل ساده رگرسیون خطی

امکان تخمین پارامترها

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}},$$
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

در نتیجه امکان محاسبه خط رگرسیون $\hat{y} = \widehat{\beta_0} + \widehat{\beta_{۱}} x$

به ازای هر x_i امکان محاسبهٔ مقدار برازش $\widehat{y_i}$ از $\widehat{y_i} = \widehat{\beta_0} + \widehat{\beta_{۱}} x_i$

$\widehat{y_i}$ مقدار پیش‌بینی شدهٔ y_i

باقیمانده یا خطا

$$\mathrm{e}_i = \widehat{y_i} - y_i$$

روش اول حل مدل ساده رگرسیون خطی

الگوریتم محاسبهٔ $\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} x$

- ورودی: (x_1, y_1)، (x_2, y_2)، ....، (x_n, y_n)
- تخمین

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}},$$
$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}$$

- به طوری که

$$s_{xx} = \sum_{i=1}^{n} (x_i - \overline{x})^2,$$
$$s_{xy} = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})$$

روش اول حل مدل ساده رگرسیون خطی

مثال $(1,3) \quad (2,4) \quad (3,8) \quad (4,9)$

حل

$$\bar{x} = \frac{1+2+3+4}{4} = 2.5,$$
$$\bar{y} = \frac{3+4+8+9}{4} = 6,$$
$$s_{xx} = (1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2 = 5,$$
$$s_{xy} = (1-2.5)(3-6) + (2-2.5)(4-6) + (3-2.5)(8-6) + (4-2.5)(9-6) = 11$$

داریم

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}} = \frac{11}{5} = 2.2$$
$$\hat{\beta}_0 = 6 - (2.2)(2.5) = 0.5$$

روش اول حل مدل ساده رگرسیون خطی

$$(1,3) \quad (2,4) \quad (3,8) \quad (4,9)$$

مثال

$$\hat{y}_i = 0.5 + 2.2x_i,$$

ادامهٔ حل

$$\hat{y}_1 = 2.7, \quad \hat{y}_2 = 4.9, \quad \hat{y}_3 = 7.1, \quad \hat{y}_4 = 9.3$$

$$\begin{aligned}
e_1 &= y_1 - \hat{y}_1 = 3 - 2.7 = 0.3, \\
e_2 &= y_2 - \hat{y}_2 = 4 - 4.9 = -0.9, \\
e_3 &= y_3 - \hat{y}_3 = 8 - 7.1 = 0.9, \\
e_4 &= y_4 - \hat{y}_4 = 9 - 9.3 = -0.3
\end{aligned}$$

$$, e_1 + e_2 + e_3 + e_4 = 0$$

روش دوم حل مدل ساده رگرسیون خطی

روش کمترین مربعات

ورودی: $(x_۱,y_۱)$، $(x_۲,y_۲)$، ....، (x_n,y_n)

خطا برابر با $e_i = y_i - \hat{y_i} = y_i - \beta_0 - \beta_۱ x$

بدست آوردن جمع مربع خطاها

$$g(\beta_0,\beta_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

بهترین برازش برابر با کمینهٔ تابع بالا

$$\frac{\partial g}{\partial \beta_0} = \sum_{i=1}^{n} 2(-1)(y_i - \beta_0 - \beta_1 x_i) = 0,$$

$$\frac{\partial g}{\partial \beta_1} = \sum_{i=1}^{n} 2(-x_i)(y_i - \beta_0 - \beta_1 x_i) = 0$$

روش دوم حل مدل ساده رگرسیون خطی

حل معادلات

$$\frac{\partial g}{\partial \beta_0} = \sum_{i=1}^{n} 2(-1)(y_i - \beta_0 - \beta_1 x_i) = 0,$$

$$\frac{\partial g}{\partial \beta_1} = \sum_{i=1}^{n} 2(-x_i)(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}},$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

$$s_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

$$s_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

گسترش

رگرسیون خطی چندگانه

$$y = \beta_0 + \beta_1 x + \beta_2 z + \cdots + \beta_k w + \epsilon$$

معنای خطی بودن

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

مشکلات

- بیش‌برازش، ناهم‌وراریانسی، هم‌خطی چندگانه

منابع تهیه این درس

[پینسکی]

[راس]

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

مرزبان و سانچز

پیشرو-نیک

[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