شناسایی آماری الگو

مروری بر آمار و احتمال

محمدجواد فدائی‌اسلام

فصل اول: آمار توصیفی

شاخص‌های گرایش مرکزی:

۱-میانگین

الف-میانگین حسابی

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

ب- میانگین هندسی

$$G = \sqrt[n]{x_1 x_2 \ \ldots x_n} = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

۲- میانه

۳- نما (مد)

شاخص‌های گرایش مرکزی:

۲- میانه

ویژگی‌ها:

الف- میانه مشاهدات را به دو بخش مساوی تقسیم می‌کند.

ب- منحصر به فرد است.

ج- تحت تأثیر داده‌های پرت قرار نمی‌گیرد.

د- محاسبه آن ساده است.

۳- نما (مد)

نمای یک مجموعه، عددی است که در آن مجموعه بیش از بقیه تکرار شده باشد.

شاخص های پراکندگی

۱- دامنه

$$R = x_{max} - x_{min}$$

۲-واریانس

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

۳-انحراف معیار

۴-متغیرهای استاندارد

۲- ویژگی‌های واریانس

۱- واریانس عدد ثابت C برابر با صفر است.

۲- اگر مقدار ثابت a را به مشاهدات اضافه یا از آنها کم کنیم واریانس تغییر نمی‌کند.

۳- اگر مشاهدات در مقدار ثابت K ضرب یا برآن تقسیم شود واریانس جدید از ضرب یا تقسیم واریانس قدیم درK^2 بدست می آید

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

۳- انحراف معیار

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

۴- متغیرهای استاندارد

$$z_i = \frac{x_i - \bar{x}}{s}$$

- **ویژگی‌های متغیرهای استاندارد:**

۱- میانگین متغیرهای استاندارد برابر صفر است.

۲- واریانس متغیرهای استاندارد برابر با ۱ است.

۳- متغیرهای استاندارد فاقد واحد اندازه‌گیری هستند.

۴- مقدار Z_i می تواند، منفی، صفر یا مثبت باشد.

فصل دوم: احتمال

فضای نمونه

- مجموعه ای از همه برآمدهای ممکن یک تجربه تصادفی را فضای نمونه می‌گویند. و آن را با علامت S نمایش می دهند.
  - یک سکه را آنقدر پرتاب می کنیم تا رو (H) ظاهر شود. فضای نمونه را بنویسید.

$$S=\{H,TH,TTH,\ldots\}$$

که S گسسته و نامتناهی شمارا است

- **پیشامد**
  - هر زیرمجموعه از فضای نمونه را یک پیشامد گویند.

احتمال

مفهوم کلاسیک:

$$P(A) = \frac{\text{تعداد حالات مساعد}}{\text{تعداد حالات کل}} = \frac{n(A)}{n(S)}$$

مفهوم فراوانی: احتمال یک پیشامد، برابر با نسبت دفعاتی است که پیشامدهای از یک نوع، در تکرار زیاد رخ خواهند داد، احتمال به مفهوم فراوانی تلقی می‌شود.

تابع احتمال

تابعی را که به هر پیشامد، عددی در بازه (۱،۰) نسبت دهد و در سه اصل زیر صدق کند تابع احتمال گویند.

<u>**اصل اول:**</u> احتمال هر پیشامد بزرگتر یا مساوی صفر است.

$$P(A) \geq 0, \forall A \subset S$$

<u>**اصل دوم:**</u> احتمال فضای نمونه S برابر با ۱ است.

$$P(S) = 1$$

<u>**اصل سوم:**</u> در صورت استقلال پیشامدها داریم.

$$P[A_1 \cup A_2 \dots] = P(A_1) + P(A_2) + \cdots$$

قوانین احتمال

قضیه: احتمال مجموعه تهی برابر صفر است.

$$P(\phi) = 0$$

قضیه: اگر A' متمم پیشامد A باشد آنگاه

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

قضیه: اگر $A \subset B$ باشد آنگاه $P(A) \leq P(B)$

قضیه: اگر A یک پیشامد باشد آنگاه $0 \leq P(A) \leq 1$

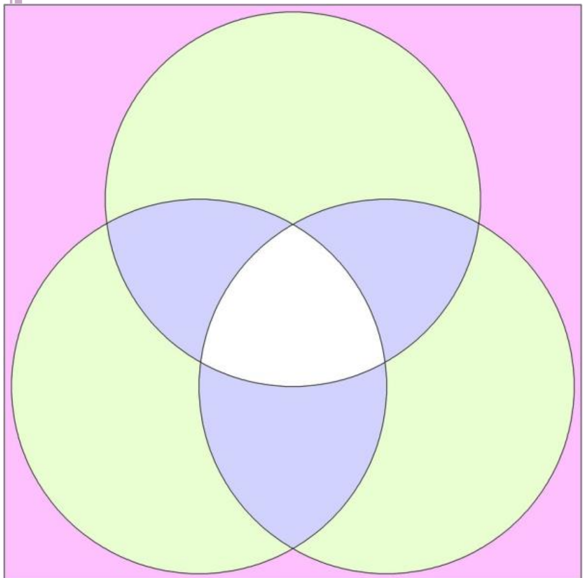
احتمال ...

قضیه: اگر A،B دو پیشامد دلخواه در S باشند آنگاه:

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

قضیه: اگر A، B و C پیشامدهای دلخواه در S باشند آنگاه:

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



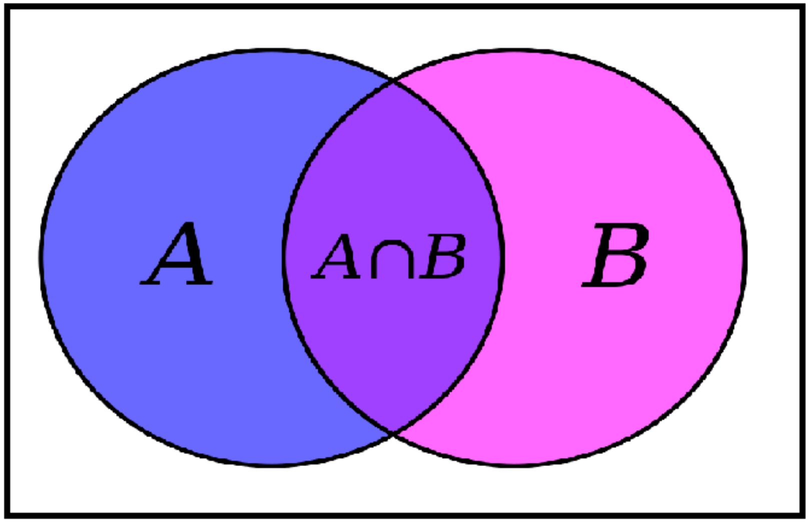
احتمال شرطی

احتمال شرطی پیشامد A به شرط وقوع پیشامد B به صورت زیر تعریف می‌شود

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0$$

وقوع B به این معنی است که فضای حالت S به B کاهش یافته است و

وقوع A به $A \cap B$ تقلیل یافته است.



احتمال شرطی

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

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

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

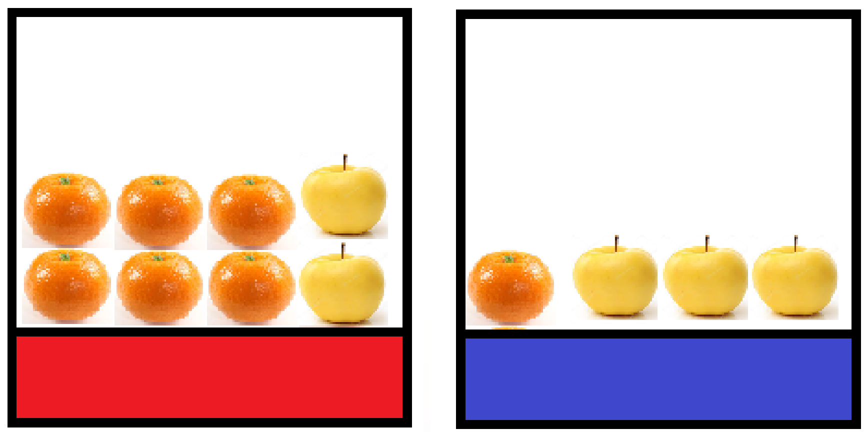
احتمال شرطی - مثال

دو جعبه قرمز و آبی را در نظر بگیرید که شامل تعدادی سیب و پرتقال به صورت زیر است. جعبه‌ها با احتمال داده شده به صورت تصادفی انتخاب می‌شود.

$$P(B=r)=\frac{4}{10}, P(B=b)=\frac{6}{10},$$

مطلوب است محاسبه احتمالات زیر:

$$P(F=a|B=r)=1/4$$
$$P(F=o|B=r)=3/4$$
$$P(F=a|B=b)=3/4$$
$$P(F=o|B=b)=1/4$$



دو پیشامد مستقل

دو پیشامد A و B را مستقل گوییم، اگر رخ داد یکی تأثیری در دیگری نداشته باشد

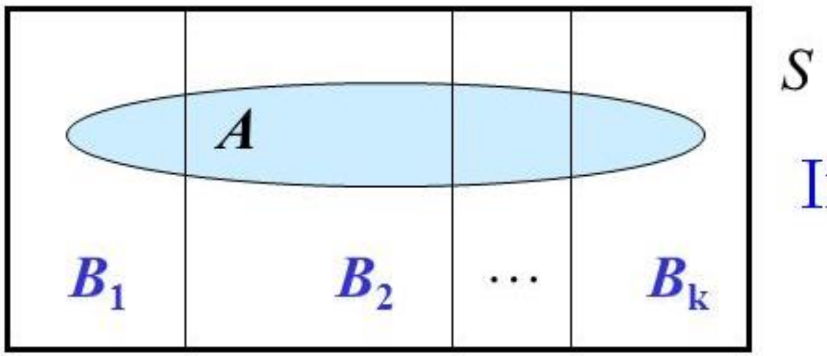
یعنی $P(A|B) = P(A)$ و $P(B|A) = P(B)$.

بنابراین A و B مستقل اند اگر:

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

مثال: روآمدن در پرتاب سکه با عدد ۶ آمدن در پرتاب تاس

قضیه احتمال کلی



If $P(B_i) \neq 0$, then

$$P(A \cap B_i) = P(A | B_i)P(B_i)$$

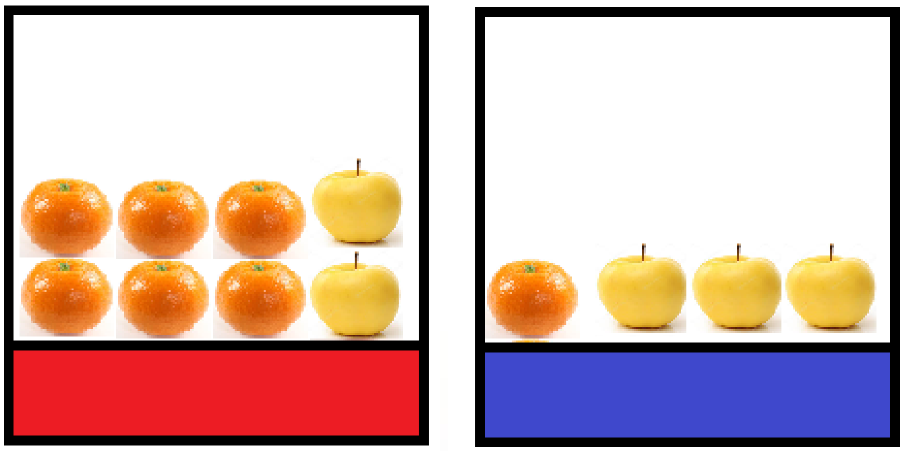
Hence, for the partition $\{B_1, B_2, \ldots, B_k\}$ of the sample space S, we have $P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)$

or equivalently,

$$\boxed{P(A) = P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \cdots + P(A | B_k)P(B_k).}$$

احتمال کلی - مثال

- $P(F = o) = P(F = o|B = r)P(B = r)$
  $+ P(F = o|B = b)P(B = b) = \frac{3}{4} \times \frac{4}{10} + \frac{1}{4} \times \frac{6}{10} = \frac{9}{20}$
- $P(F = a) = 1 - \frac{9}{20} = \frac{11}{20}$



قضیه بیز

اگر پیشامدهای A_1 ، A_2 ، ...، A_k دو به دو مجزا و $S = \bigcup_{i=1}^{k} A_i$ باشد احتمال شرطی هریک از A_iها به شرط اتفاق پیشامد A از S برابر با:

$$P(A_i|A) = \frac{P(A_i \cap A)}{P(A)}$$

$$P(A_i|A) = \frac{P(A_i \cap A)}{\sum_{i=1}^{k} P(A|A_i)P(A_i)}$$

قضیه بیز یکی از مهمترین روابط مورد استفاده در شناسایی الگو است.

قضیه بیز برای طبقه‌بندی

$$P(w_j \mid x) = \frac{P(x \mid w_j)P(w_j)}{\sum_{k=1}^{N} P(x \mid w_k)P(w_k)} = \frac{P(x \mid w_j)P(w_j)}{P(x)}$$

- که x بردار ویژگی و w_j کلاس j ام است.
- هر کلاسی که $P\left(w_j \middle| x\right)$ برای آن حداکثر شود انتخاب می‌شود.
- هریک از اجزا در رابطه بیز دارای نام خاصی است:
- $P(w_j)$ احتمال پیشین کلاس w_j - prior probability
- $P\left(w_j \middle| x\right)$ احتمال پسین کلاس w_j به هنگام x – posterior probability
- $P\left(x \middle| w_j\right)$ احتمال شرطی

قضیه بیز – مثال

- $P(B=r|F=o)=\frac{P(F=o|B=r)P(B=r)}{P(F=o)}=\frac{3}{4}\times\frac{4}{10}\times\frac{20}{9}=\frac{2}{3}$
- $P(B=b|F=o)=1-\frac{2}{3}=\frac{1}{3}$