

Session 11

Fundamentals of Probability

Applied Mathematics for AI | Hobot Academy | Zahra Amini



Telegram & Instagram & YouTube: @hobotacademy & LinkedIn: @zahraamini-ai

www.hobotacademy.com

احتمال شاخه‌ای از ریاضیات است که عدم قطعیت را مدل‌سازی می‌کند. در یادگیری ماشین، بسیاری از مدل‌ها و الگوریتم‌ها برای تصمیم‌گیری و استنتاج به احتمال تکیه دارند. به عنوان مثال، شبکه‌های بیزین و مدل‌های مارکوف پنهان از اصول احتمال برای پیش‌بینی و استنتاج استفاده می‌کنند.

احتمال چیست؟

ارتباط بین احتمال و مدل‌های یادگیری آماری

یادگیری آماری روشی برای تخمین توابع از داده‌ها است که بر پایه نظریه احتمال بنا شده است. در بسیاری از الگوریتم‌های یادگیری، مانند رگرسیون لجستیک، احتمال برای مدل‌سازی توزیع خروجی‌ها به کار می‌رود.

تعريف فضای نمونه
Sample Space

مجموعه تمامی نتایج ممکن از یک آزمایش تصادفی.

در پرتاب یک تاس، فضای نمونه $\{1, 2, 3, 4, 5, 6\}$ است.

زیرمجموعه‌ای از فضای نمونه که وقوع آن مورد توجه است.

Event

محاسبه احتمال
یک رویداد

Classical Probability: $P(A) = \frac{|A|}{|S|}$

احتمال یک رویداد A در یک فضای نمونه متناهی، به صورت زیر تعریف می‌شود:

فرض کنید یک تاس سالم (با ۶ وجه) پرتاب می‌کنیم. فضای نمونه عبارت است از:

$$S = \{1, 2, 3, 4, 5, 6\}$$

حال می‌خواهیم احتمال رویداد A را که نشان‌دهنده‌ی ظاهر شدن عددی زوج است، محاسبه کنیم. عددی زوج در فضای نمونه:

$$A = \{2, 4, 6\}$$

پس تعداد حالات مطلوب $3 = |A|$ و تعداد کل حالات ممکن $6 = |S|$ ، بنابراین:

$$P(A) = \frac{3}{6} = \frac{1}{2} = 0.5$$

تمرین ۱: اگر یک کارت از یک دسته کارت استاندارد ۵۲ تایی بکشیم، احتمال آنکه کارت کشیده شده یک دل (Hearts) باشد چیست؟

تمرین ۲: فرض کنید در یک کلاس ۳۰ دانش‌آموز وجود دارد که ۱۸ نفر آن‌ها دختر و ۱۲ نفر آن‌ها پسر هستند. اگر یک دانش‌آموز به طور تصادفی انتخاب شود، احتمال آنکه دختر باشد چقدر است؟

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

قاعده جمع:

Multiplication Rule: $P(A \cap B) = P(A)P(B|A)$

قاعده ضرب:

Complement Rule: $P(A^c) = 1 - P(A)$

قاعده مکمل:

فرض کنید دو سکه سالم را به طور همزمان پرتاب می‌کنیم. فضای نمونه عبارت است از:

$$S = \{HH, HT, TH, TT\}$$

$$A: \text{سکه اول رو باشد} \rightarrow P(A) = 1/2$$

$$B: \text{سکه دوم رو باشد} \rightarrow P(B) = 1/2$$

از آنجایی که پرتاب سکه‌ها مستقل از یکدیگر است، احتمال اینکه هر دو سکه رو باشند:

$$P(A \cap B) = P(A)P(B) = 1/2 \times 1/2 = 1/4$$

فرض کنید یک کلاس ۳۰ دانشآموز دارد و یکی از دانشآموزان علی است. احتمال اینکه اگر یک دانشآموز به‌طور تصادفی انتخاب شود علی باشد برابر است با:

$$P(A) = 1/30$$

$$P(A^c) = 1 - P(A) = 1 - 1/30 = 29/30 \approx 0.97$$

پس احتمال اینکه فرد انتخاب شده علی نباشد برابر است با:

فرض کنید یک تاس سالم پرتاب می‌شود. فضای نمونه عبارت است از:

$$S = \{1, 2, 3, 4, 5, 6\}$$

دو رویداد زیر را تعریف می‌کنیم:
 $A: \text{اظاهر شدن عددی زوج} \rightarrow P(A) = 3/6 = 0.5$

$B: \{3, 2, 1\} \rightarrow \text{اظاهر شدن عددی کمتر از } 4 \rightarrow P(B) = 3/6 = 0.5$

$A \cap B: \text{عددی که هم زوج باشد و هم کمتر از } 4 \rightarrow P(A \cap B) = 1/6$

حال طبق قاعده جمع:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.5 - 1/6 \approx 0.83$$

مثال

احتمال شرطی بیانگر احتمال وقوع یک رویداد، مشروط به رخدادن رویداد دیگر است:

$$\text{Conditional Probability: } P(A|B) = \frac{P(A \cap B)}{P(B)}$$

از قانون ضرب می‌توان برای محاسبه احتمال وقوع چندین رویداد به طور همزمان استفاده کرد:

$$\text{Bayes' Theorem: } P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$\text{Independence: } P(A \cap B) = P(A)P(B)$$

قضیه بیز ارتباط بین احتمال شرطی دو رویداد را مشخص می‌کند:

فرض کنید که ۵٪ از ایمیل‌ها اسپم هستند و ۸۰٪ از اسپم‌ها حاوی کلمه «برنده» باشند، در حالی که فقط ۱۰٪ از ایمیل‌های غیر اسپم

این کلمه را دارند. احتمال اینکه ایمیلی که شامل «برنده» است، اسپم باشد را می‌توان با قضیه بیز محاسبه کرد:

$$P(\text{Spam}|\text{Winner}) = \frac{P(\text{Winner}|\text{Spam})P(\text{Spam})}{P(\text{Winner})}$$

Given Data $\rightarrow P(\text{Spam}) = 0.05, \quad P(\text{Winner}|\text{Spam}) = 0.8, \quad P(\text{Winner}|\neg\text{Spam}) = 0.1, \quad P(\neg\text{Spam}) = 1 - P(\text{Spam}) = 0.95$

Step 1: Compute $P(\text{Winner}) \rightarrow P(\text{Winner}) = P(\text{Winner}|\text{Spam})P(\text{Spam}) + P(\text{Winner}|\neg\text{Spam})P(\neg\text{Spam})$

$$P(\text{Winner}) = (0.8 \times 0.05) + (0.1 \times 0.95) \rightarrow P(\text{Winner}) = 0.04 + 0.095 = 0.135$$

Step 2: Compute $P(\text{Spam}|\text{Winner}) \rightarrow P(\text{Spam}|\text{Winner}) = \frac{P(\text{Winner}|\text{Spam})P(\text{Spam})}{P(\text{Winner})}$

$$P(\text{Spam}|\text{Winner}) = \frac{(0.8 \times 0.05)}{0.135} \rightarrow P(\text{Spam}|\text{Winner}) = \frac{0.04}{0.135} = 0.296$$

Final Result $\rightarrow P(\text{Spam}|\text{Winner}) = 0.296$