

Session 11

Fundamentals of Probability

Applied Mathematics for AI | Hobot Academy | Zahra Amini



Hobot Academy

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

www.hobotacademy.com

احتمال چیست؟

ارتباط بین احتمال و مدل‌های یادگیری آماری

تعریف فضای نمونه
Sample Space

رویداد **Event**

محاسبه احتمال
یک رویداد

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. به‌عنوان مثال، شبکه‌های بیزین و مدل‌های مارکوف پنهان از اصول احتمال برای پیش‌بینی و استنتاج استفاده می‌کنند.

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.

مجموعه تمامی نتایج ممکن از یک آزمایش تصادفی. در پرتاب یک سکه، فضای نمونه $\{H, T\}$ است.

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

زیرمجموعه‌ای از فضای نمونه که وقوع آن مورد توجه است.

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

احتمال یک رویداد A در یک فضای نمونه‌ی متناهی، به‌صورت 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.

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

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

قوانین احتمال

مثال

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

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

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

B : 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 تاس سالم پرتاب می‌شود. فضای نمونه عبارت است از:

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

دو رویداد زیر را تعریف می‌کنیم:

A : ظاهر شدن عددی زوج $\rightarrow P(A) = 3/6 = 0.5$

B : $\{3, 2, 1\} = 4$ ظاهر شدن عددی کمتر از ۴ $\rightarrow P(B) = 3/6 = 0.5$

$A \cap B$: عددی که هم زوج باشد و هم کمتر از ۴ $\{2\} \rightarrow P(A \cap B) = 1/6$

حال طبق قاعده جمع:

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

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:

$P(A) = 1/30$

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

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

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

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

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

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

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

قضیه بیز ارتباط بین احتمال شرطی دو رویداد را مشخص می کند:

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