جلسه چهارم آمار و احتمال مهندسی "دکتر شریفی زارچی - دانشگاه شریف"

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

احتمال شرطی (Conditional probability)

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

قاعده بیز (Bayes's Rule)

قانون احتمال کل

اگر Ω افراز شده باشد به $B_1, B_2, \dots$ آنگاه

$$P(A) = \sum P(A|B_i)\, P(B_i)$$

اثبات

$$\Omega = \bigcup_i B_i \;,\; A = A\cap\Omega \Rightarrow A = A \cap \left(\bigcup_i B_i\right) = \bigcup_i (A\cap B_i)$$

$$\Rightarrow P(A) = P\left(\bigcup_i (A\cap B_i)\right) = \sum P(A\cap B_i) = \sum P(A|B_i)\, P(B_i) \quad \boxtimes$$

مثال دو سکه با احتمالات H، T طبق جدول داریم. یک سکه را انتخاب کرده و پرتاب می‌کنیم.

	H	T
سکه نوع 1	1/2	1/2
سکه نوع 2	1/4	3/4

a) احتمال 3 بار شیر آمدن در 4 بار پرتاب سکه‌ی نوع اول چقدر است.

جدول برای اینکه بعد از انتخاب شدن هر سکه و 4 بار پرتاب برای هر حالت چه احتمالاتی وجود دارد. →

	0	1	2	3	4
اول	$(\frac{1}{2})^4$	$\binom{4}{1}\times\frac{1}{2}\times(\frac{1}{2})^3$	$\binom{4}{2}\times(\frac{1}{2})^4$	$\binom{4}{3}\times(\frac{1}{2})^3\times\frac{1}{2}$	$(\frac{1}{2})^4$
دوم	$(\frac{1}{4})^4$	$\binom{4}{1}(\frac{1}{4})^3\times\frac{3}{4}$	$\binom{4}{2}\times(\frac{1}{4})^2\times(\frac{3}{4})^2$	$\binom{4}{3}(\frac{3}{4})^3\times\frac{1}{4}$	$(\frac{3}{4})^4$

b) یک سکه را تصادفی انتخاب و 4 بار پرتاب کردیم. 3 بار شیر آمده. چقدر احتمال دارد که سکه از نوع 1 باشد؟

A: سکه‌ی انتخاب شده، از نوع اول باشد.

B: رویداد اینکه سکه‌ی پرتاب شده، از 4 بار، 3 بار شیر بیاید.

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

$P(B|A) = \frac{1}{4}$

$P(B|A^c) = (\frac{3}{4})^3$

$P(A|B) = ?$

$$P(A|B) = \frac{P(B|A)\, P(A)}{\underbrace{P(B|A)\, P_A + P(B|A^c)\, P(A^c)}_{P(B)}} = \frac{16}{43}$$

c) چقدر احتمال دارد که در مسئله b سکه از نوع 2 باشد؟

$$P(A^c|B) = 1 - P(A|B) = 1 - \frac{16}{43}$$

قانون بیز به حالت کلی

اگر Ω به $B_1, B_2, \ldots$ افراز شده باشه آن‌گاه:

$$P(B_i|A) = \frac{P(A|B_i)\,P(B_i)}{\underbrace{\sum_j P(A|B_j)\,P(B_j)}_{P(A)}}$$

کاربرد در پردازش تصویر (Machine Learning)

		B_0	B_1	B_2	....	B_9
		0	1	2		9
/	1	0.01				
7	2	0.001				
Ɔ	3	1				

یک سری ویژگی در دست خط و نوشتار اعداد رو در نظر بگیریم مثل خط صاف یا شکستگی و یا انحنا. یک مجموعه بزرگ از دست نوشته های اعداد جمع آوری می کنیم. برای تک تک داده های مجموعه بدست میاریم که این داده چه عددی هست و کدوم ویژگی ها رو داره. در انتها درصد می گیریم که مثلا چند درصد از 0 های مجموعه داده ی ورودی ویژگی 1 رو که خط صاف باشه داشتن. چند درصد ویژگی 2 که شکستگی باشه رو داشتن و ... . در انتها جدول ما برای اعداد 0 تا 9 در میاد :)

برای هر عدد مثلا 0 چند مدل نوشتار با توجه به ویژگی ها داریم که شامل و یا فاقد چه ویژگی هستند. و حالا در بین تمام داده های 0 که داریم. چند درصد از 0 ها به هر کدام از دسته بندی ها قرار می گیرند؟ این مدل جدول رو برای تمام B_i ها بدست می آوریم.

B_0

1	X	✓	X	X	✓	✓	X	✓
2	X	X	✓	X	✓	X	✓	✓
3	X	X	X	✓	X	✓	✓	✓
	0.01	0.001	0.2	0.75	.	- -	-	0.15

$P(B_i|A)$: تصویر ورودی ما A هست. چقدر احتمال داره که خالقش عدد i باشه؟

$P(A|B_i)$: ویژگی های تصویر A رو مثل جدول دوم استخراج می کنیم. ویژگی های A $\boxed{\begin{matrix}\checkmark\\ X\\ X\end{matrix}}$ باشه این موقع مقدار $P(A|B_i)$ رو از جدول دوم B_i می خوایم.

$P(B_i)$: مقدار دفعات تکرار عدد i به کل دیتاست (به صورت کلی تعداد تکرارهای اعداد در دیتاست ها یکسان نیست)

در نهایت برای هر عدد i مقدار $P(B_i|A)$ رو بدست میاریم. و i که مقدار $P(B_i|A)$ رو ماکزیمم کرده به عنوان خروجی تصویر A برمی گردونیم.

و ما الان واقعا یک الگوریتم ML داریم. :)

نکته: برای تشخیص / ۷ د به صورت خالص از الگوریتم استفاده می‌شود.

کنجکاوی: جنسیت بچه اول و بچه دوم خانواده از هم مستقلند؟

فرزند اول	فرزند دوم	a)	b)
د	د	1/4	1/10
د	پ	1/4	2/10
پ	د	1/4	3/10
پ	پ	1/4	4/10

A : دختر بودن فرزند اول

B : دختر بودن فرزند دوم

استقلال

زمانی گوییم که دو واقعه مستقلند که یکی از دو شرط زیر برقرار باشد.

(1) $\begin{cases} P(B|A) = P(B) \\ \text{and} \\ P(A|B) = P(A) \end{cases}$ (2) $\begin{cases} P(A) = 0 \\ \text{or} \\ P(B) = 0 \end{cases}$

از نظر شهودی دونستن اطلاعات هرکدام تأثیری روی احتمالات دیگری نداشته باشه.

پاسخ a

$$P(B|A) = \frac{P(B\cap A)}{P(A)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}, \quad P(B) = \frac{1}{2}$$

$$P(A|B) = \frac{P(B\cap A)}{P(B)} = \frac{1/4}{1/2} = 1/2, \quad P(A) = \frac{1}{2}$$

پس مستقلند.

پاسخ b

$$P(B|A) = \frac{P(B\cap A)}{P(A)} = \frac{\frac{1}{10}}{\frac{6}{10}} = \frac{1}{6}, \quad P(B) = \frac{4}{10}$$

$$P(A|B) = \frac{P(B\cap A)}{P(B)} = \frac{\frac{1}{10}}{\frac{4}{10}} = \frac{1}{4}, \quad P(A) = \frac{3}{10}$$

پس A و B وابسته‌اند.

قضیه

A , B مستقلند اگر و فقط اگر $P(A\cap B) = P(A)\cdot P(B)$

اثبات

$\Rightarrow$ فرض A,B مستقلند حکم $P(A\cap B)=P(A)\cdot P(B)$

if $P(A)$ or $P(B) = 0 \Rightarrow P(A\cap B) = P(A)\cdot P(B) = 0$ ✓

$$\text{if } \begin{cases} P(B|A) = P(B) \\ P(A|B) = P(A) \end{cases} \xrightarrow[\text{احتمال شرطی}]{\text{تعریف}} P(A|B) = \frac{P(A\cap B)}{P(B)} = P(A) \Rightarrow P(A\cap B) = P(A)\times P(B)$$ ✓

$\Leftarrow$ فرض $P(A\cap B) = P(A)\times P(B)$ حکم: A, B مستقلند

$$\begin{cases} P(A|B) = \dfrac{P(A\cap B)}{P(B)} \overset{\text{فرض}}{=} \dfrac{P(A)\times P(B)}{P(B)} = P(A) \\[2ex] P(B|A) = \dfrac{P(B\cap A)}{P(A)} \overset{\text{فرض}}{=} \dfrac{P(A)\times P(B)}{P(A)} = P(B) \end{cases}$$ ✓

نکته

زمانی دو واقعه مستقلند که هر خانه در جدول روبه‌رو برابر با حاصل ضرب Marginal probability حاشیه باشد.

برای مثال در جدول روبه‌رو A, B مستقل نیستند.

	B	B^c		
A	$1/10$	$2/10$	$3/10$	$\frac{1}{10}+\frac{2}{10}$
A^c	$3/10$	$4/10$	$7/10$	$\frac{3}{10}+\frac{4}{10}$
	$4/10$	$6/10$		
	$\frac{1}{10}+\frac{3}{10}$	$\frac{4}{10}+\frac{2}{10}$		

Marginal probabilities $= \Sigma$ سطرها یا ستون‌ها