PAI. Probability Tutorial

Anastasia Makarova

ETH Zürich

27.09.2019

Probability

Today:

- Conditional Probability
- Sum Rule and Product Rule
- Conditional independence
- Conditional and marginal distributions

Conditional Probability

- Let X and Y - random variable with p(x) and p(y)
- In general, $p(x, y) \neq p(x)p(y)$
  If X and Y are independent, $p(x, y) = p(x)p(y)$.
- Conditional probability $p(x|y) = \frac{p(x,y)}{p(y)}$
  Meaning: how the fact $Y = y$ affects the distribution of X.
  Note, that $\int p(x|y)dx \equiv 1$, but $\int p(x|y)dy$ is not essentially 1, because for y it is not a density function, but likelihood function.

Sum Rule

- All the operations over probabilities are based on sum rule and product rule.
- Sum rule: Let $X_1, \ldots, X_k$ – are mutually excluding, one of which always happens. Then:

$$P(X_i \cup X_j) = P(X_i) + P(X_j) \qquad \sum_{i=1}^{k} P(X_i) = 1$$

- In discrete case: $\forall Y : \sum_{i=1}^{k} P(X_i|Y) = 1$, then

$$\sum_{i=1}^{k} \frac{P(Y|X_i)P(X_i)}{P(Y)} = 1 \qquad P(Y) = \sum_{i=1}^{k} P(Y|X_i)P(X_i)$$

- In continuous case:

$$p(y) = \int p(y, x)dx = \int p(y|x)p(x)dx$$

Product Rule

- Product rule: any joint distribution can be represented as a product

$$p(x, y) = p(x|y)p(y) \qquad P(X, Y) = P(X|Y)P(Y)$$

- Equivalent form: p(x|y)p(y) = p(x,y) = p(y|x)p(x)
- For multidimensional joint distributions:

$$p(x_1, \ldots, x_n) = p(x_1|x_2, \ldots, x_n) \ldots p(x_{n-1}|x_n)p(x_n)$$

- For a pdf of n variables, i.e., $p(x_1, x_2, \ldots x_n)$, how many different factorizations exist?
- If the variables are all independent, how many different factorizations exist?

Conditional and marginal distributions

- We are interested in distribution of x_a of a multivariate random variable (x_a, x_b). How to do so?
- We know values $p(x_b)$: conditional distribution

$$p(x_a|x_b) = \frac{p(x_a, x_b)}{p(x_b)}$$

- We don't know values $p(x_b)$: marginal distribution

$$p(x_a) = \int p(x_a, x_b) dx_b$$

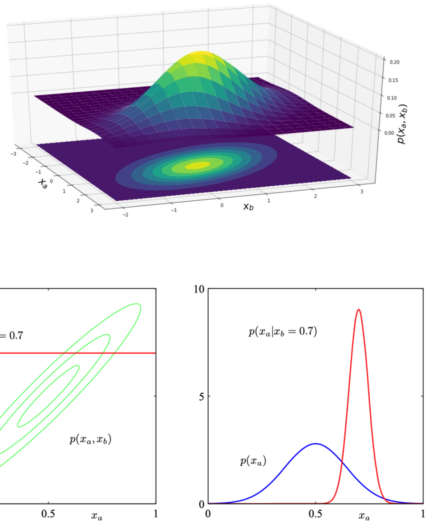

Figure: Upper: Joint distribution $p(x_a, x_b)$ in 3D.
Lower: Joint distribution $p(x_a, x_b)$, conditional $p(x_a|x_b)$, marginal $p(x_a)$
(Image: Bayesian methods, HSE, Vetrov)

Bayes Rule

- $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
- 1% of women at age forty who participate in routine screening have breast cancer. 80% of women with breast cancer will get positive mammography. 9.6% of women without breast cancer will also get positive mammography.
  A women in this age group had a positive mammography in a routine screening. What is the probability that she has breast cancer?

Bayes Rule

Data

- $P(BC = T) = 1\%,\ P(BC = F) = 99\%.$
- $P(+|BC = T) = 80\%,\ P(+|BC = F) = 9.6\%.$

$$
\begin{aligned}
P(BC = T|+) &= \frac{P(BC = T, +)}{P(+)} \\
&= \frac{P(BC = T, +)}{\sum_{BC=\{T,F\}} P(BC, +)} \\
&= \frac{P(+|BC = T)P(BC = T)}{\sum_{BC=\{T,F\}} P(+|BC)P(BC)} \\
&= \frac{80\% \times 1\%}{80\% \times 1\% + 9.6\% \times 99\%} = 7.76\%
\end{aligned}
$$

Conditional Independence

Prove or disprove (by counterexample):

(a) $X \perp Y | Z \Rightarrow X \perp Y$

False. Assume that: $P(X=1) = 0.1$, $P(Y=1) = 0.5$, $P(Z=1) = 0.5$, and:

$P(\cdot \mid Z=z)$	$Z=0$	$Z=1$
$X=1$	0.4	0.3
$Y=1$	0.6	0.2

$$P(X,Y) = \sum_Z P(X,Y,Z) = \sum_Z P(X|Z)P(Y|Z)P(Z)$$

$$P(X,Y) = P(X|Y)P(Y)$$

$$P(X=1, Y=1) = 0.4 \times 0.6 \times 0.5 + 0.3 \times 0.2 \times 0.5 = 0.15$$

$$P(X=1|Y=1) = P(X=1, Y=1)/P(Y=1) = 0.3 > P(X=1)$$

Conditional Independence

Prove or disprove (by counterexample):

(b) $(X \perp Y|Z) \& (X \perp Z|Y) \Rightarrow X \perp (Y, Z)$

$$
\begin{aligned}
P(X, Y, Z) &= P(X, Y|Z)P(Z) = P(X|Z)P(Y|Z)P(Z) \\
&= P(X|Z)P(Y, Z) \\
&= P(X, Z|Y)P(Y) = P(X|Y)P(Z|Y)P(Y) \\
&= P(X|Y)P(Y, Z) \\
&\Rightarrow P(X|Z) = P(X|Y)
\end{aligned}
$$

$$
\begin{gathered}
P(X|Z)P(Z)P(Y) = P(X, Z)P(Y) = \\
P(X|Y)P(Y)P(Z) = P(X, Y)P(Z)
\end{gathered}
$$

$$
\underbrace{\sum_z P(X, Z = z)}_{=P(X)} P(Y) = P(X, Y) \underbrace{\sum_z P(Z = z)}_{=1} \Rightarrow X \perp Y
$$

Conditional Independence

- Does conditional independence implies independence?
- Does independence implies conditional independence?

Conditional Independence

- Does conditional independence imply independence? (proved in the tutorial)
- Does independence imply conditional independence? (hw exercise)