Lecture 3 - parametric models

Thursday, September 12, 2024 2:33 PM

today: · prob. review
· parametric models

Bayes rule

$\rightsquigarrow$ invert conditioning

$$p(x|y) = \frac{p(y|x)\,p(x)}{p(y)}$$

by def.: $p(y|x) = \frac{p(x,y)}{p(x)}$

$$\Rightarrow p(x,y) = p(y|x)\,p(x) = p(x|y)\,p(y)$$

"product rule"

example of conditioning: $P\{$having lung cancer | tumor measurement = "_"$\}$

Chain rule: → successive application of product rule

$$p(x_1,\dots,x_n) = p(x_n|x_{1:n-1})\,\underbrace{p(x_{1:n-1})}$$

$$= \quad '' \quad p(x_{n-1}|x_{1:n-2})\,p(x_{1:n-2}) \dots$$

$$\boxed{p(x_1,\dots,x_n) = \prod_{i=1}^{n} \underbrace{p(x_i|x_{1:i-1})}}$$

always true

[can also choose any $n!$ permutations]

Convention: $1{:}0 = \emptyset$

$$p(x_1|x_\emptyset) = p(x_1)$$

can be simplified by making conditional independence assumptions

[for a directed graphical model]

$$p(x_i|x_{1:i-1}) = p(x_i|x_{\pi_i})$$

parents of i

conditional independence:

X is conditionally independent of Y given Z notation $X \perp\!\!\!\perp Y \mid Z$

$\Leftrightarrow p(x,y|z) = p(x|z)\,p(y|z)$ $\forall x \in \Omega_X$, $\forall y \in \Omega_Y$, $\forall z \in \Omega_Z$ s.t. $p(z) \neq 0$

"factorization view"

exercise to the reader: prove that $X \perp\!\!\!\perp Y \mid Z \Rightarrow p(x|y,z) = p(x|z)$

[conditional analog of $X \perp\!\!\!\perp Y \Rightarrow p(x|y) = p(x)$]

[conditional analog of $X \perp\!\!\!\perp Y \Rightarrow p(x|y) = p(x)$]

example: Z = indicator whether mother carries a genetic disease

X = " daughter 1 has disease $\qquad X \perp\!\!\!\perp Y \mid Z$

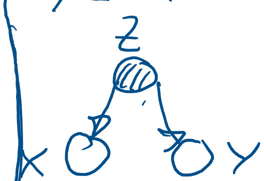
Y = " " 2 " "

aside:

$X \perp\!\!\!\perp Y$

$X \not\perp\!\!\!\perp Y \mid Z$

$\hookrightarrow p(x,y,z) = p(z|x,y)\, p(y)\, p(x)$



here, X is not "marginally indep" of Y

(i.e. $X \not\perp\!\!\!\perp Y$) $\exists\, x,y$ s.t. $p(x,y) \neq p(x)p(y)$

$\hookrightarrow p(x,y,z) = p(x|z)\, p(y|z)\, p(z)$

example of pairwise indep R.V. that are not mutually independent

X = coin flip as $\{0,1\}$ $\qquad$ define $Z \triangleq X \text{ xor } Y$

Y = another indep. coin flip

$Z \perp\!\!\!\perp Y$

$Z \perp\!\!\!\perp X$

$X \perp\!\!\!\perp Y$

$Z \not\perp\!\!\!\perp (X, Y)$

other notation:

expectation/mean of a R.V.

$$\mathbb{E}[X] \triangleq \sum_{x \in \Omega_X} x\, p(x) \quad \text{(for discrete R.V.)}$$

$$\int x\, p(x)\, dx \quad \text{(" cts. R.V.)}$$

$\mathbb{E}[\cdot]$ is a linear operator $\Rightarrow \mathbb{E}[\underset{\text{scalar}}{a} X + Y] = a\,\mathbb{E}[X] + \mathbb{E}[Y]$ $\quad$ R.V.'s

Variance

$$\mathrm{Var}[X] \triangleq \mathbb{E}[(X - \mathbb{E}[X])^2] \quad \text{"measure of dispersion"}$$

$$= \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$\sigma = \sqrt{\mathrm{Var}[X]}$

$\mathbb{E}[X]$

cumulative distribution fct. CDF of X: $F_X(t) = P\{X \leq t\}$ $\left(= \int_{-\infty}^{t} p(x)\,dx \right.$

$$\frac{d}{dt} F_X(t) \Big|_{t=x} = p(x) \rightarrow \text{(for a cts. R.V.)}$$

parametric models:

↳ a family of distributions

$$\mathcal{P}_{\Theta} = \{ \underbrace{p(\cdot\,;\theta)}_{} \mid \theta \in \Theta \}$$

possible pmf's / pdf's

depend on parameter θ

abuse of notation $\{ p(x;\theta) \mid \theta \in \Theta \}$

"better" notation $\{ p(X;\theta) \mid \theta \in \Theta \}$

from context

"support" of family i.e. Ω_X

is usually fixed for all $\theta \in \Theta$

(but not always)

examples: $\Omega_X = \{0,1\}$ for a coin flip

$\Omega_X = [0,+\infty$ for a gamma dist. family

notation: $X \sim \text{Bern}(\theta)$

"R.V. X is distributed as a Bernoulli dist. with parameter θ"

$$p(x;\theta) = \text{Bern}(x;\theta) \triangleq \theta^{x}(1-\theta)^{1-x} \qquad x \in \{0,1\}$$

variable for pmf

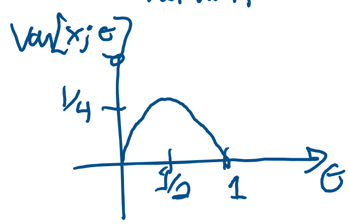
parameter of dist.

Bernoulli: prob of a coin flip $\Omega_X = \{0,1\}$ $\Theta = [0,1]$

$p(x=1) = \theta$ Bayesian notation $P\{X=1 \mid \theta\} = \theta$

$\mathbb{E}[X] = \theta$

$\text{Var}[X] = \theta(1-\theta)$



another example:

$$p(x;(\mu,\sigma^2)) = N(x;(\mu,\sigma^2)) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

or

$N(x \mid \mu, \sigma^2)$

<u>binomial distribution</u>:

model n indep. coin flips

↳ sum of n indep $\text{Bern}(\theta)$ R.V.

let $X_i \overset{\text{iid}}{\sim} \text{Bern}(\theta)$ (mutually) independent and identically distributed

implicitly talking $X_1, \dots, X_n$ $\quad (X_i)_{i=1}^{n} \overset{\text{iid}}{\sim} \dots$

let $X \triangleq \sum_{i=1}^{n} X_i$ then we have that $X \sim \text{Bin}(n, \theta)$

"binomial dist. with parameter n & θ"

$\Omega_X = \{0, 1, \dots, n\}$

pmf: $p(x; \theta) = \underbrace{\binom{n}{x}}_{\substack{\text{\# of ways to choose } x \\ \text{elements out of } n \text{ possibilities}}} \theta^{x} (1-\theta)^{n-x}$

"n choose x"

derivation:

$\text{Bern}(x_i; \theta) = \theta^{x_i} (1-\theta)^{1-x_i}$ $\triangleq 0$

$p(x_1, \dots, x_n) \underset{\text{by indep.}}{=} \prod_{i=1}^{n} \text{Bern}(x_i; \theta) = \theta^{\sum_i x_i} (1-\theta)^{n - \sum_i x_i}$

$\binom{n}{x} \triangleq \dfrac{n!}{x!\,(n-x)!}$

mean: $X = \sum_i X_i$

$\mathbb{E}[X] = \sum_i \mathbb{E}[X_i] = n\theta$

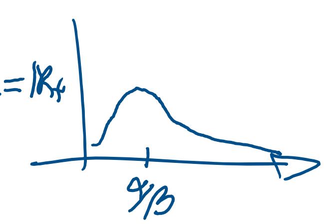
similarly, $\text{Var}\left(\sum_i X_i\right) \overset{\text{indep.}}{=} \sum_i \text{Var}(X_i) = n\theta(1-\theta)$

<u>other distribution</u>: $\text{Poisson}(\underset{\text{mean}}{\lambda})$ $\quad \Omega_X = \{0, 1, 2, \dots\} = \mathbb{N}$ [count data]

Gaussian in 1 $\quad N(\underset{\text{mean}}{\mu}, \underset{\text{variance}}{\sigma^2})$ $\quad \Omega_X = \mathbb{R}$

gamma dist. $\quad \Gamma(\underset{\text{shape}}{\alpha}, \underset{\text{inverse scale}}{\beta})$ a.k.a. rate $\quad \Omega_X = \mathbb{R}_+$

mean: $\dfrac{\alpha}{\beta}$ variance $\dfrac{\alpha}{\beta^2}$



other: Laplace, Cauchy, exponential dist. , beta dist. is a Dirichlet on 2 elements

$\hookrightarrow \Gamma(1,\beta)$ $\hookrightarrow \Omega_X = [0,1]$