LECTURE 11: 09/11 — Law of total prob., Bayes Formula & applications.

THM' 10.11: Let $A_1, \ldots, A_n$ be events in Ω $\ni$ $\mathbb{P}(A_1 \cap \cdots \cap A_n) > 0$.

Then $\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1 \mid A_2 \cap \cdots \cap A_n) \times \mathbb{P}(A_2 \mid A_3 \cap \cdots \cap A_n)$

$\times \cdots \times \mathbb{P}(A_{n-1} \mid A_n) \times \mathbb{P}(A_n)$.

(Chain rule for condnl. prob.)

THM 10.12: Let $\Omega = \bigsqcup_{i=1}^{n} A_i$ & $\mathbb{P}(A_i) > 0$ $\forall i$.

Then $\forall B \subseteq \Omega$ $\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \mid A_i)\,\mathbb{P}(A_i)$.

[Law of Total prob]

Proof: $n=2$. Assume $\Omega = A_1 \sqcup A_2$.

$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \mathbb{P}(B \cap A_2)$ (finite additivity & $B = B\cap A_1 \sqcup B \cap A_2$.)

$= \mathbb{P}(B \mid A_1)\mathbb{P}(A_1) + \mathbb{P}(B \mid A_2)\mathbb{P}(A_2)$ (defn of $\mathbb{P}(\cdot \mid A_i)$)

Ex: Prove the general case. ∎

Medical test: If a patient has a disease, the test returns positive with 95% prob.

Eg No 1 If a patient is healthy, - - - with 10% prob.

5% of population has the disease.

What is the prob. someone is healthy given they test positive?

SOLUTION: $\Omega = \{H, D\} \times \{P, N\}$ $\quad \bar{H} = H \times \{P, N\}, \bar{D}$

Given information: $\mathbb{P}(\bar{H}) = 0.95$. $\quad \mathbb{P}(\bar{D}) = 0.05$

$\bar{P} = \{H, D\} \times P$; $\mathbb{P}(\bar{P} \mid \bar{H}) = 0.10$ $\quad \mathbb{P}(\bar{P} \mid \bar{D}) = 0.95$.

+ve

Question $\quad \mathbb{P}(\bar{H} \mid \bar{P}) = \text{Prob. (healthy} \mid \text{test is positive)}$

$$\mathbb{P}(\bar{H} \mid \bar{P}) = \frac{\mathbb{P}(\bar{H} \cap \bar{P})}{\mathbb{P}(\bar{P})} =$$

[$\bar{H}$ – Person is healthy; $\bar{D}$ – ; $\bar{P}$ – Person tests positive, $\bar{N}$ –]

$$= \frac{\mathbb{P}(\bar{P}\,|\,\bar{H})\,\mathbb{P}(\bar{H})}{\mathbb{P}(\bar{P}\,|\,\bar{H})\,\mathbb{P}(\bar{H}) + \mathbb{P}(\bar{P}\,|\,\bar{D})\,\mathbb{P}(\bar{D})}$$ (using defn & LTP)

[We can use LTP b'cos $\Omega = \bar{H} \sqcup \bar{D}$ & $\mathbb{P}(\bar{H}), \mathbb{P}(\bar{D}) > 0$]

$$= \frac{0.10 \times 0.95}{0.10 \times 0.95 + 0.95 \times 0.05} = \frac{2}{3}$$

Test is bad. Think why? Most people are healthy but also test works badly on 10% healthy!

Ex. What should $\mathbb{P}(\bar{P}\,|\,\bar{D})$, $\mathbb{P}(\bar{P}\,|\,\bar{H})$ be so as to make error < 0.01

More serious error is $\mathbb{P}(\bar{D}\,|\,\bar{N})$?

Find a "cheap test" to minimize $\underbrace{\mathbb{P}(\bar{D}\,|\,\bar{N})}_{\text{very small}}$ even if $\underbrace{\mathbb{P}(\bar{H}\,|\,\bar{P}) \gtrsim 0.1}_{\text{reasonable}}$.

THM 1.2 (Bayes Formula): Let $\Omega = \bigcup_{i=1}^{n} A_i$, $P(A_i) > 0 \ \forall i$.

Then $$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}(B \mid A_i)\, \mathbb{P}(A_i)}{\sum_{j=1}^{n} \mathbb{P}(B \mid A_j)\, \mathbb{P}(A_j)}$$ for $\mathbb{P}(B) > 0$.

Proof:

$$\mathbb{P}(A_i \mid B) \underset{\text{(defn)}}{=} \frac{P(B \cap A_i)}{P(B)} = \frac{P(B \mid A_i)\, P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)\, P(A_j)}$$ — defn, — LTP.

∎

Eg 1.3 Consider medical test as before.

$\mathbb{P}(\bar{P} \mid \bar{H}) = 0.02$ i.e., 2% healthy test positive.

$$\mathbb{P}(\bar{D} \mid \bar{N}) = \frac{\mathbb{P}(\bar{N} \mid \bar{D})\, \mathbb{P}(\bar{D})}{\mathbb{P}(\bar{N} \mid \bar{D})\, \mathbb{P}(\bar{D}) + \mathbb{P}(\bar{N} \mid \bar{H})\, \mathbb{P}(\bar{H})}$$

diseased negative

$$= \frac{0.05 \times 0.05}{0.05 \times 0.05 + 0.98 \times 0.95} \qquad (\mathbb{P}(\overline{N} \mid \overline{D}) = 1 - \mathbb{P}(\overline{P} \mid \overline{D}))$$

$$\approx 0.0026.$$

Ideal test: $\mathbb{P}(\overline{N} \mid \overline{D}) = 0$; $\mathbb{P}(\overline{N} \mid \overline{H}) = 1.$