Lecture 1

1/8/2015

Branching process, MCMC, Bayes Melding, Bootstrap
↳ big topics to be covered in a course

Modeling
study model system to gain insight into real world
study interacting systems in mathematical modeling
↳ computational hypotheticals

- Classical Methods: Anderson & May
- Random Process Models
- Data fitting

} course will cover

Sensitivity: probability test positive given diseased
$s = P(E \mid D)$

Die Example

$\Omega = \{1, 2, 3, 4, 5, 6\}$ Sample space
Set of possible outcome

$1 \in \Omega$

$7 \notin \Omega$

Set $A = \{2, 4\}$ $A \subseteq \Omega$ A is a subset of Ω ← subset symbol

$\omega \in A \Rightarrow \omega \in \Omega$

$B = \{5\}$

Union ($\cup$) : OR

$A \cup B = \{2, 4, 5\}$

Intersection ($\cap$) : AND

$A \cap B = \{\} = \emptyset$ The Empty Set

} A and B are disjoint

b/c its one or the other happens
both don't happen
[b/c $A \cap B = \emptyset$]

Disjoint union

$P(A \cup B) = P(A) + P(B)$

$p(\Omega) = 1$

Something that happens every time (relative frequency 1)

has probability 1

Probability is the long run limit of the relative frequency

<u>Specificity</u> $f = P(\bar{E} \mid \bar{D})$

E = event test +

$\bar{E}$ = event test −

D = event diseased

$\bar{D}$ = event not diseased

$$P(E \mid D) = \frac{P(E \cap D)}{P(D)}$$

$E = E \cap D \cup E \cap \bar{D}$

By definition:

$P(E \cap D) = P(E \mid D) P(D)$

$(E \cap D) \cap (E \cap \bar{D}) = \phi$

$\therefore \quad P(E) = P(E \cap D) + P(E \cap \bar{D})$

$\quad = P(E \mid D) P(D) + P(E \mid \bar{D}) P(\bar{D})$

θ = prob test +

$\theta = s\pi + (1-f)(1-\pi)$

<u>Law of Total Probability</u>

$P(E \cap D) = P(D) P(E \mid D)$

$\quad = P(E) P(D \mid E)$

Therefore, $P(D \mid E) = \dfrac{P(D) P(E \mid D)}{P(E)}$ $\boxed{\text{Bayes Theorem}}$

Bayes Rule follows from axioms of probability
↳ Frequentists use it!

Bayesians: probability represents degree of belief

Can prove that if there are degrees of belief they obey axioms of probability & therefore are probabilities

Prior — Probability of evidence given your belief

$$P(D|E) = \frac{P(D)\,P(E|D)}{P(E)}$$

- Tells you how to update beliefs w/ evidence
- How to treat beliefs rationally
- Rational agents should converge on the same beliefs w/ overwhelming

A & B are <u>independent</u> if $P(A|B) = P(A)$

Also $P(A \cap B) = P(A)\,P(B)$ if A indep of B

<u>3-way Independence</u>:

$P(A \cap B \cap C) = P(A)\,P(B)\,P(C)$

$P(A \cap B) = P(A)\,P(B)$

⋮ + each pairwise independence

etc.

<u>Needle Reuse Example</u>: 2-State Markov Chain (Next time!)

- Rinse single needle & reuse on multiple patients

Random Variable (RV) — random number produced by a random process

Bernoulli RV — 2 outcomes, eg coin test for disease
- success
- failure

Needle stick problem continued

Success - hepatitis infection

N needle sticks on N patients

Distribution of number of successes?

# of successes is a RV X

$N = 51$

$$X = X_1 + X_2 + \ldots + X_{50} + X_{51}$$

$$P(X=x) = \binom{N}{x} p^x (1-p)^{N-x}$$

Multiplications are a result of the independence assumption

Binomial Probability Formula

HIV Infection Example

π prev of HIV, β prob of infection from a partner

$1-(1-\pi\beta)^n$ chance of infection w/ n partners

⇝ can estimate β using maximum likelihood (future lecture!)

Expectation

$$EX = E[X] = \sum_x x P(X=x)$$

Def of mean, expected value, average

Example:

x	$P(X=x)$
0	0.7
1	0.2
2	0.1

$$EX = 0\,P(X=0) + 1\,P(X=1) + 2\,P(X=2)$$

$$= 0 + 0.2 + 2 \times 0.1 = 0.4$$

1. $E[X+Y] = E[X] + E[Y]$ expected value of a sum is the sum of the expected value

Proof:

$$E[X+Y] = \sum_x \sum_y (x+y)\, P(X=x, Y=y)$$

$$= \sum_x \sum_y x\, P(X=x, Y=y) + \sum_y \sum_x y\, P(X=x, Y=y)$$

$$= \sum_x x \sum_y P(X=x, Y=y) + \sum_y y \sum_x P(X=x, Y=y)$$

$$= \sum_x x\, P(X=x) + \sum_y y\, P(Y=y)$$

$$= EX + EY$$

Example

51 Bernouilly Trials

$$X = X_1 + \cdots + X_{51}$$

$$EX = EX_1 + \cdots + EX_{51} = \overbrace{p + p + \cdots + p}^{51} = 51p$$

Generally:

$EX = Np$ expectation of a Binomial RV

Variance $\operatorname{var} X = E[(X-EX)^2]$ Definition

$$E[(X-EX)^2] = E[X^2 - 2XEX + (EX)^2]$$

$$= E[X^2] + E[-2X\,EX] + E[(EX)^2]$$

$$= E[X^2] - 2EX\, E[X] + (EX)^2$$

$$= E[X^2] - (E[X])^2 \geq 0$$

$\boxed{\text{Useful fact: } E[kX] = kEX}$

The second moment is always greater than the square of the mean.

$$EXY = \sum_x \sum_y xy\, P(X=x, Y=y)$$

Assume $X Y$ indep

$$EXY = \sum_x \sum_y xy\, P(X=x)\, P(Y=y)$$

$$= \sum_x x\, P(X=x) \sum_y P(Y=y)$$

Variance of a Sum

$$\operatorname{var}(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$$

$$= EX^2 + 2EXY + EY^2 - (EX)^2 - 2EXEY - (EY)^2$$

$$\boxed{\text{indep} \Rightarrow EXY = EXEY}$$

$$= EX^2 - (EX)^2 + EY^2 - (EY)^2$$

$$= \operatorname{var} X + \operatorname{var} Y$$

Variance of a Binomial

$$EX = p, \quad \operatorname{var} X = \sum_x x^2 P(X=x) = 0^2 P(X=0) + 1^2 P(X=1) = p$$

$$\operatorname{var}(X_1 + \ldots + X_N) = N \operatorname{var} X_1 = Np(1-p)$$

Needle Stick example

Lets say 3 infections out of 51 infections

what is estimate of p?

$$\hat{p} = \frac{3}{51}$$

[Generally $\hat{p} = \frac{X}{N}$]

$$E\hat{p} = E\left[\frac{X}{N}\right] = \frac{1}{N} EX = \frac{1}{N} Np = p$$

$\hat{p}$ is unbiased

Variance of kX

$$\begin{aligned} var[kX] &= E[(kX)^2] - (E[kX])^2 \\ &= k^2 EX^2 - k^2 (EX)^2 \\ &= k^2 \, var X \end{aligned}$$

$$\begin{aligned} var\,\hat{p} = var\left[\frac{X}{N}\right] &= \frac{1}{N^2} var X \\ &= \frac{1}{N^2} Np(1-p) \end{aligned}$$

$$var\,\hat{p} = \frac{p(1-p)}{N}$$

$$SE(\hat{p}) = \sqrt{\frac{p(1-p)}{N}}$$