08/07/19

AMS 131

Lecture 5

section review: ELISA / HIV study

$+$ = ELISA says HIV +

$-$ = ELISA says HIV −

prevalence: $P(A) = 0.004$

sensitivity: $P(+ \mid A) = 0.96$

specificity: $P(- \mid \text{not } A) = 0.97$

So what is $P(A \mid +)$?

$\underset{\text{truth}}{A} \quad \underset{\text{data}}{+}$

$P(A \mid B)$ not always $= P(B \mid A)$

$P(A \text{ and } B) = P(A) \cdot P(B \mid A) = P(B)\, P(A \mid B)$

$$P(B \mid A) = \frac{P(B)\,P(A \mid B)}{P(A)} \qquad P(\text{cause} \mid \text{effect}) = \frac{P(C)\,P(E \mid C)}{P(E)}$$

$$P(\text{truth} \mid \text{data}) = \frac{P(T)\,P(D \mid T)}{P(D)}$$

$\text{prior} \xrightarrow[\text{data}]{} \text{post.}$

$$P(\text{posterior}) = \frac{P(\text{prior})\,P(\text{data info})}{P(\text{normalizing})}$$

$$P(A \mid +) = \frac{P(A) \cdot P(+ \mid A)}{P(+)} = \frac{\text{prevalence} \cdot \text{sensitivity}}{\text{partitions}}$$

$\underset{\text{truth}}{A} \quad \underset{\text{data}}{+}$

		TRUTH		
		A	not A	
DATA	+	384	2988	3372
	−	16	96612	96628
		400	99600	100,000

$P(A \mid +) = \frac{384}{3372} = 11\%$

$P(A^c \mid -) = \frac{96612}{96628} = 99.98\%$

	A	A^c
+	true pos.	false pos.
−	false neg	true neg.

← ELISA says you have HIV when you really don't (false pos.)

← ELISA says you don't have, when you do (false neg)

$$\left[\frac{P(B \mid A)}{P(B^c \mid A)}\right] = \left[\frac{P(B)}{P(B^c)}\right]\left[\frac{P(A \mid B)}{P(A \mid B^c)}\right]$$

posterior odds in favor of B | A = prior odds in favor of B · Bayes factor, likelihood in favor of B

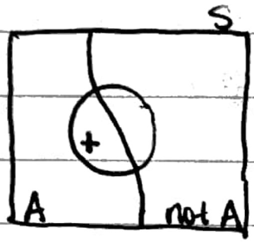
Lecture 5 (cont.) AMS 131

Method 3: Evaluate $P(+)$ if we know the truth

$\left.\begin{matrix} A \\ \text{not } A \end{matrix}\right]$ truth $\left.\begin{matrix} + \\ - \end{matrix}\right]$ data

in $P(+ \mid A)$, $P(+)$ is easier to calculate

"extending the conversation": to compute probability about the data bring in the truth. (partition)



$P(+) = P[(+ \text{ and } A) \text{ or } (+ \text{ and not } A)]$

$(A, \text{not } A)$ = partition

$P(+ \text{ and } A) + P(+ \text{ and not } A)$
$= P(A)P(+ \mid A) + P(A^c)P(+ \mid A^c)$
$= (0.004)(0.96) + (0.996)(0.03) = 0.03372$

Law of Total Probability

Bayes Theorem $$P(B \mid A) = \frac{P(B)P(A \mid B)}{\sum_{i=1}^{k} P(C_i)P(A \mid C_i)}$$ ← partitions

B_i: truths background info $P(B_i)$

A: data after info $P(B_i \mid A)$

$$P(B_i \mid A) = \frac{P(B_i)P(A \mid B_i)}{\sum_{j=1}^{k} P(B_j)P(A \mid B_j)}$$

↑ partition the truth

$P(A \mid B_i)$ represents the likelihood that data set A would be if B_i was the actual state

$P(A)$ not relying on B_i = normalizing constant

$$\sum_{j=1}^{k} P(B_j)P(A \mid B_j) = P(A)$$. denominator

$$\text{posterior} = \frac{(\text{prior})(\text{likelihood})}{\text{normalization}}$$

Lecture 5 (cont.)

Random Variable Distribution: given a sample space S in an experiment $\mathcal{E}$, a random variable (RV) is a function from the non-weird collection C to the real number line $\mathbb{R}$.

T-S study: $Y(TNNTN) = 2$ Y ignores the order.
$Y(NNNTT) = 2$

↳ similar to a 'function machine'

$P(Y=y) = P(\{s: Y(s) = y\})$

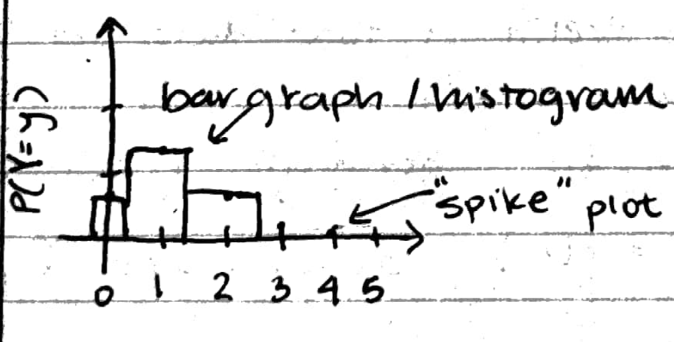
T/F propos. set

$P(Y=1) = P(\{s \in S: Y(s) = 1\})$
$= P(\{TNNNN, NTNNN \ldots NNNNT\})$

y	$P(Y=y)$
0	0.237
1	0.396
2	0.264
3	0.088
4	0.015
5	0.001

Y = random variable
y = possible value of Y

R.V. can be shown by its possible values and their probabilities.



A distribution of a R.V. Y is the collection of all probabilities $P(Y \in A)$ for all sets A in non-weird $C_{\mathbb{R}}$

Lecture 5 (cont.) AMS 131

A random variable has a discrete distribution, or Y is a discrete R.V. if the set of possible values Y is finite or at most countable infinite. R.V. for which the set of possible values is uncountable are call continuous R.V.

$$X = \begin{cases} 1 & \text{if } Y > 0 \\ 0 & \text{otherwise} \end{cases} \qquad Y = \#\text{ of T-S babies}$$

$\{0, 1\}$ dichotomous, binary

Imagine weighing things where you can specify the number of sig figs. With a 1 pound (1 lb) package:

<u>possible weights</u>

16 oz

16.0 oz

15.99 oz

15.993 oz

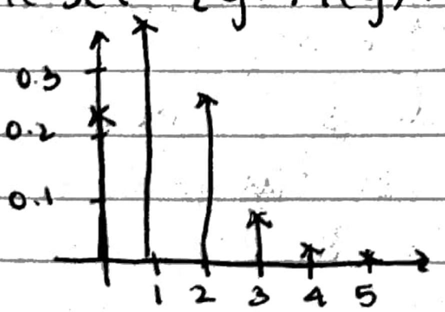
15.9928 oz

If there is no conceptual limit to the number of sig figs., you could get a R.V. Y = the actual weight, modeled as continuous, having $(0, \infty)$ the positive part of $\mathbb{R}$

Conceptually continuous: all numbers around 16 are possible = continuous. (i.e. no limit in principle to precision of measurement).

Probability (mass) function: f_Y that keeps track of the probabilities associated with Y $f_Y(y) = P(Y = y)$

The set $\{y : f_Y(y) > 0\}$ is called the support of Y



PF (PMF) of T-S study

AMS 131

Lecture 5 (cont.)

A. R.V Y that only takes on the values $\{0,1\}$ is said to have a Bernoulli distribution with parameter p - written Bernoulli(p)

if $\underset{(PMF)}{f_Y(y)} = P(Y=y) = p^y(1-p)^{1-y} = \begin{cases} p & \text{for } y=1 \\ 1-p & \text{for } y=0 \end{cases}$ 0 else

similar to indicator $I(A) \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false} \end{cases}$ $\qquad I_{\{0,1\}}(y)$ ↑ support

set ↓

$I_A(y) \begin{cases} 1 & \text{if } y \in A \\ 0 & \text{else} \end{cases}$

Freq.

$Y \sim$ Bernoulli (p)

↑ "is distributed as"

Bayesian

$(Y|p) \sim$ Bernoulli (p)

Powerball Lottery $\{1, 2, \ldots 69\}$

5 balls drawn W_i = # on iTH ball

$P(W_i = w_1) = \begin{cases} \frac{1}{69} & w_1 = 1, 2, \ldots 69 \\ 0 & \text{otherwise} \end{cases}$

same for W_2, W_3 if nothing is known about the previous draws.

For any two integers $a \leq b$ a R.V. Y that is equally likely to be any of the values $\{a, a+1 \ldots b\}$ has the uniform distribution

Uniform $\{a, b\}$

its pf is $f_Y(y) = P(Y=y) = \begin{cases} \frac{1}{b-a+1} & y = a \ldots b \\ 0 & \text{else} \end{cases}$

$Y \sim$ uniform $\{a, b\}$ $\leftrightarrow$ Y chosen at random from $\{a, a+1, \ldots b\}$

Lecture 5 (cont.) AMS 131

In random trials preformed, recorded as a success or failure. If each trial is independent and the chance p of success is constant, then

Y = # of success for binomial distribution

$$pf = f_Y(y) = P(Y=y) = \begin{cases} \binom{n}{y} p^y (1-p)^{(n-y)} & \text{for } y = 0, 1 \ldots n \\ 0 & \text{else} \end{cases}$$

$Y \sim$ binomial (n, p)

Let $B_i = \begin{cases} 1 & \text{if trial success} \\ 0 & \text{failure} \end{cases}$ for $i, \ldots, n$ $\quad \underline{\text{sum}(B_i) = Y}$

$B_i \overset{IID}{\sim}$ Bernoulli(p), all B_i independent

$X_i \overset{IID}{\sim} f_{X_i}(x_i)$ all R.V $X_1, X_2 \ldots$

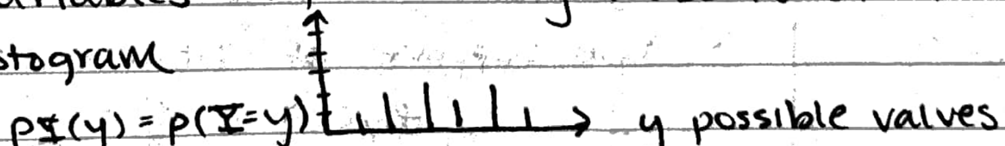
are independent and identically distributed with pf $f_{X_i}(x_i)$

$$\left(Y = \sum_{i=1}^{n} B_i\right) \sim \text{Binomial}(n, p)$$

distribution of the sum of bunch of IIDs R.Vs

discrete random variables → probability mass function

↳ spike spot / histogram



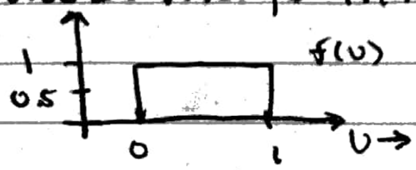
single percision: 7 sig figs. $3.141592|653589$

floating point 3 round off error

errors can accumulate / propagate

Uniform $\{0, 0.1 \ldots 0.9\}$ and $\{0, 0.01 \ldots 0.99\}$

- would merge into one block



discrete uniform dist. to continuous uniform dist.

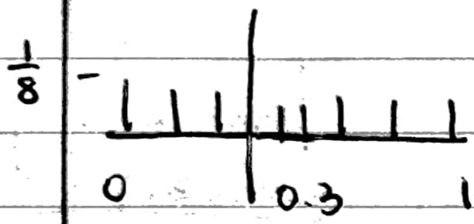
Uniform (0, 1)

Smooth, continuous function (analogue of discrete pf)

$$f_{U}(u) = \begin{cases} 1 & \text{for } 0 \le u \le 1 \\ 0 & \text{else} \end{cases}$$

probability density function (vs. mass function)

analogue of summation is integration



PMF: $P(U \le 0.3)$

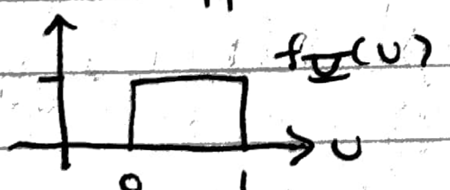
$P(U = 0 \text{ or } \frac{1}{8} \text{ or } \frac{2}{8})$

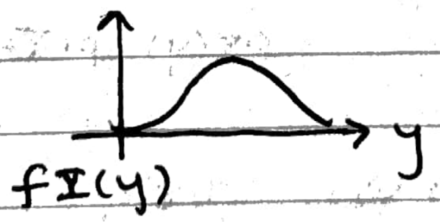
$$\sum_{u=0,\frac{1}{8},\frac{2}{8}} f_{U}(u)$$

A random variable (Y has a continuous distribution) $\leftrightarrow$ (Y is a continuous R.V.)

if there exists a continuous non-negative function f_Y defined on $\mathbb{R}$ for every interval $[a,b]$, $P(a \le Y \le b) = \int_a^b f_Y(y)\,dy$

If Y is a continuous R.V, f_Y is a probability density function of Y. The set $\{y : f_Y(y) > 0\}$ is the support of (the dist. of) Y





normal / Gaussian

PMF	PDF
$f_Y(y) \ge 0$	$f_Y(y) \ge 0$
$\sum f_Y(y) = 1$	$\int_{-\infty}^{\infty} f_Y(y)\,dy = 1$

Lecture 5 (cont.) AMS 131

$P(a \le Y \le b) = \int_a^b f_Y(y)\,dy$

continuous function, Y cont. R.V.

$\int_a^a f_Y(y)\,dy = 0 \quad \therefore P(Y = a) = 0$ for all $a \in \mathbb{R}$

single point = singleton $\{a\}$ otherwise $\int \to \infty$

probability for continuous random variables
$\leftrightarrow$ area under PDF

$P(Y=y)=0$ for $-\infty \le y \le \infty$, singletons have to have zero probability due to integration