Amanda Chapin

AMS 7 10/13/09

EXPERIMENTAL DESIGN & PROBABILITY

- HW#2 due Tuesday Oct 20th
- Lab#2 due Friday Oct 23rd
- Reading: through ch8 + ch9
- should be able to finish through prob#3, HW#2 today

(L-86) • a design is valid (unbiased) if: you repeat the design many times & the average of the results would give you the truth

How to improove accuracy of a CRD
(a) get genetically pure strain (sibs)
(b) choose 60 litters at random
(c) choose 2 from each litter & assign T/C

1. completely randomized design
2. MATCHED PAIRS DESIGN
   - CRD held constant

T_1	Enriched	RANDOMIZED BLOCKS DESIGN →
T_2	Normal	same design as above w/ 3+
T_3	Deprived	variables instead of 2 (MATCHED PAIRS)

• matched pairs special case of randomized blocks (blocks=2)

(L-88) • 100 ppl w/ imsomnia given placebo C, & drug T

$$\begin{array}{c|c|c} & T & C \\ 1 & & \\ 2 & & \\ \vdots & & \\ 100 & & \end{array}$$

LONGITUDINAL DATA: example of REPEATED - MEASURES DESIGN

• same person over time eliminates CRD better than Matched Pair

(AC)

(L-89)
- CROSS SECTIONAL → many different subjects at one moment in time (photograph)
- LONGITUDINAL → one subject studied over a long period of time (movie)

3.1 MEANING OF PROBABILITY

(L-97)
- Tay Sachs (T-S) genetic disease; fatal reduction of Hex-A enzyme
- Carrier → 1/2 normal enzyme production
- - healthy, but carriers are able to pass recessive gene on to progeny

- man & woman are both carrier & want to have big family w/ 5 children

Q: what is the probability that 1+ of their children will be born w/ T-S?

Pascal, Fermat (1650); gambling & mathmatics

(L-98)
① FREQUENTIST (RELATIVE FREQUENCY) → repeatable, identical conditions, completely independent of each other
- dice, roulett, coin toss

P(A) of an event; A = LONG RUN RELATIVE FREQ
- if I roll a dice x times, how many times would you expect to roll A?

② BAYESIAN APPROACH is more accurate, but the math is too difficult for this class

AC

$A = \{ 1^+$ T-S babies of 5 children from (Hh) parents $\}$

L-99 • work out individual prob. of each child w/ Punnet Square

		Dad H	h
Mom	H	HH	Hh
	h	Hh	hh

1 : 2 : 1

HH = normal (100% Hex-A)
Hh = carrier (50% Hex-A)
hh = T-S (0% Hex-A)

• EQUALITY LIKELY MODEL (ELM) ⇒ You can enumerate (all the ways the repeatable phenomenon can occur) in a way that all outcomes are equally likely, then for event A:

$$\boxed{P(A) = \frac{\#\text{ of outcomes favorable to A}}{\text{total }\#\text{ of possible outcomes}}}$$

Ex: one draw (Y) $\boxed{\text{at random}}$ from (1, 2, 9)
$P(Y=9)$: $1/3 = 33\%$
L-100 $P(Y \text{ is odd})$: $2/3 = 67\%$ $\boxed{\text{ELM}}$ each # has a 33% chance

• ELM & T-S
P(normal) = 25% (1/4)
P(carrier) = 50% (2/4)
P(T-S) = 25% (1/4

({exactly 1 T-S baby} $\boxed{\text{OR}}$ {exactly 2} $\boxed{\text{OR}}$ {3} $\boxed{\text{OR}}$ {4} $\boxed{\text{OR}}$ {5}

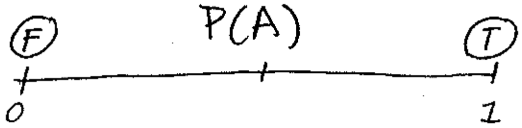
• Probability (A) vs Probability NOT (A)
P(0 T-S kids) = P(not T-S 1st $\boxed{\text{AND}}$ not 2nd $\boxed{\text{AND}}$ not 3rd $\boxed{\text{AND}}$ not 4th $\boxed{\text{AND}}$ not 5th

• use VENN DIAGRAMS when working w/ $\boxed{\text{AND}}$, $\boxed{\text{OR}}$, $\boxed{\text{NOT}}$

- if A is certain, prob A = 1 or 100%
- if A is impossible, prob A = 0 or 0%

(L-102)

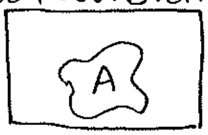
P(A)



$0 \leq P(A) \leq 1$

NOT

ALL POSSIBILITIES

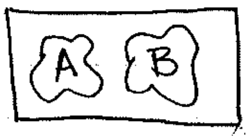


P(A) =  ← Area of blob A
100%/1

(L-103) $P(A) + P(\text{not } A) = 1$

so $P(A) = 1 - P(\text{not } A)$ *

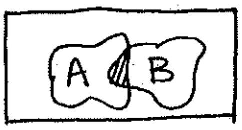
OR



(L-104)

- $P(A \text{ or } B) = P(A) + P(B)$
  - special case w/o overlap = $\boxed{\text{MUTUALLY EXCLUSIVE}}$
- $P(A \text{ or } B) = P(A) + P(B) - P(A \,\&\, B)$

General Addition Rule for $\boxed{\text{OR}}$



▨ = $P(A \,\&\, B)$

AND

$$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \xrightarrow[\text{random}]{\text{2 at}} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$p(y_1 = 2 \;\&\; y_2 = 2)$?

- INDEPENDENT IDENTICALLY DIST SAMPLING ⇒ sampling w/ replacement (IID)
- SIMPLE RANDOM SAMPLE ⇒ w/o replacement (SRS)

$\boxed{\text{IID}}$ → w/ replacement

$$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \longrightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

y_1 \ y_2	1	2	3
1	(1,1)	(1,2)	(1,3)
2	(2,1)	(2,2)	(2,3)
3	(3,1)	(3,2)	(3,3)

$P(y_1 = 2) = \frac{1}{3}$ or $\frac{3}{9}$

$P(y_2 = 2) = \frac{1}{3}$ or $\frac{3}{9}$

$P(y_1 = 2 \;\&\; y_2 = 2) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$

- JOINT DIST OF y_1 & y_2

- Theory: Prob of A & B = Prob(A) × Prob(B)

SRS → w/o replacement

$$\begin{bmatrix}1\\2\\9\end{bmatrix} \longrightarrow \begin{bmatrix}Y1\\Y2\end{bmatrix}$$

Y1 \ Y2	1	2	3
1	~~(1,1)~~	(1,2)	(1,3)
2	(1,2)	~~(2,2)~~	(2,3)
3	(3,3)	(3,2)	~~(3,3)~~

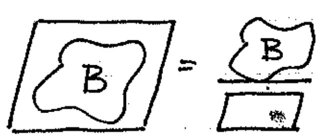
$P(Y1=2 \; \& \; Y2=2) = \frac{0}{6} = 0$
$P(Y1=2) = 1/3 = 2/6$
$P(Y2=2) = 1/3 = 2/6 \rightarrow$ Think: 2nd draw ($Y2$) depends on the first draw ($Y1$)

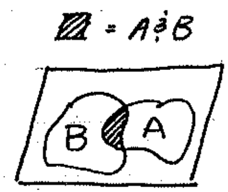
$$P(Y1=2 \; \& \; Y2=2) = 0 \neq P(Y1=2) \cdot P(Y2=2)$$

- CONDITIONAL PROBABILITY ⇒ P(B given A) = $P(B/A)$
  $= P(A \& B) / P(A)$

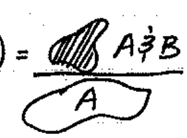
(Bays 1720)

P(B given A) = ?





P(B given A) =



- multiply by P(A) to get GENERAL PRODUCT RULE for $\boxed{\text{AND}}$

$$\boxed{\begin{aligned} P(A \& B) &= P(A) \cdot P(B/A) \\ &= P(B) \cdot P(A \& B) \end{aligned}} \text{ either order}$$

- if SRS: knowlege of first draw ($Y1$) will help predict ($Y2$)
- if IID: knowlege of first draw does not help you predict the second draw

- A & B are INDEPENDENT if knowledge of A does not predict B: second draw has same prob of first
- A & B are DEPENDENT if knowledge of A helps predict B:

AC

Theory Revision: if A & B are INDEPENDENT:

$$P(A \& B) = P(A) \cdot P(B)$$

if A & B are DEPENDENT:

$$P(B \text{ given } A) = \frac{P(B \& A)}{P(A)}$$