Julia Larson

8/13 Expiramental Design cont.

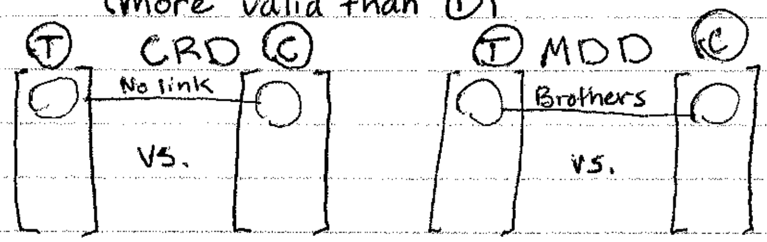
• A design is valid if it is unbiased. Meaning if you were to repeat the design many times and average the results on average the truth would emerge.

Exp. Rat Expirament

To reduce bias you must hold any PCF's constant many ways are possible to do this.

① Completely Randomized Design (valid but could be better)

② Mutual pairs design (pairs in which PCF's are held constant (more valid than ①)



Other designs

• Randomized Block Design: similar to MDD but with more subjects instedd of pair, groups exp. 3 rats from a litter T_1, T_2, T_3, matched pairs is a special case of randomized block design, with block size 2

• Longitudinal design: measure something (people) and 2 different points in time or more than 2. exp. Insomnia drug, measure people before taking drug or placebo, then after taking the drug or placebo.

• Cross section design: opposite of a longitudinal design, also called a snapshot, measures many different subjects at a single moment in time (cheaper and easier than a longitudinal design.

Probability

- The meaning of probability has 2 approaches

• Frequentist: (relative frequency) this approaches is restrict to things that are inheritely repeatable under identical conditions, with each repetition is logically independent. the probability P(A) of the event A is regarded as long run relative frequency with which A would occur in the repetition. Invented by Pascal and Fermat in 1650 to help rich people gamble better paid by king.

- Bayes Approach (1720's + 1750's): Here A can be any (True/False) proposition you want P(A) is a numerical measure of the weight of evidence in favor of the statement that A is true

T-S Disease Tay-Sachs: caused by lack of (Hex A), kill by age 4, Carriers are healthy but have only 50% the amount of (Hex A) noncarriers have 100% Hex A

Healthy HH (normal) genes 100% (Hex A)
Carrier Hh (one normal, one not normal) genes 50% (Hex A)
T-S Baby hh (two not normal) genes 0% (Hex A)

Case: a family wants to have 5 kids whats the chance of T-S babies.

Punnets Square

Father's Genes

Mother's Genes

	H	h
H	HH	Hh
h	Hh	hh

Technically this does not actually work since they will either have a T-S baby or not and they can't repeat having a baby. So we pretend lots of families wanting 5 kids are being measured.

- 2 Phenotypes: presence or abscence of the disease
- 3 Phenotypes: amount of Hex A in the blood 0%, 50% and 100%.

To help this family use the equally likely model
ELM: enumerate all the ways the repeatable phenomenon turns out, in such a way that all these possible outcomes are equally likely.

Julia Larson

$$P(A) = \frac{\text{Number of outcomes favorable to } A}{\text{Total number of possible outcomes}}$$

exp. Pop. Sample

$$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \xrightarrow[\text{Random}]{\text{at}} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

ELM

Probability of 2 being $Y_1 = 1/3$

Probability of an odd number being $Y_1 = 2/3$

Apply ELM to T-S Study: we want to know:

P (1 or more T-S kids)

P (exactly 1 T-S kids or exactly 2 T-S kids or ... or exactly 5 T-S kids)

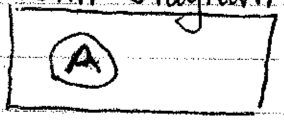
P (A or B) = P(A) ? P(B)

P (1 or more T-S kids) opposite =

P(0 T-S kids) = P (not T-S on 1st and not T-S on 2nd .... not T-S on 5th)

$P(A \text{ and } B) \overset{?}{=} P(A) \;?\; P(B)$

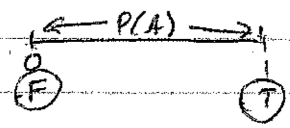
Venn Diagram (all probabilities)

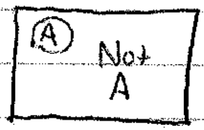


$$P(A) = \frac{\text{(A)} : \text{Area of blob (A)}}{\boxed{\text{Area of box}} : 1 \text{ or } 100\%}$$

So: If A is ceirtein P(A) = 1 = 100%

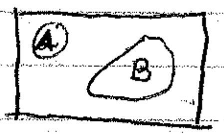
If A is impossible P(A) = 0 = 0%





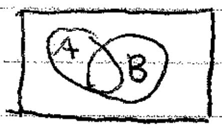
P(A) + P(Not A) = 1

so $\boxed{P(A) = 1 - P(\text{Not } A)}$ - useful equation



No overlap = Mutually exclusive

P (A or B) = P(A) + P(B)



P (A or B) = P(A) + P(B), no its too big counted the overlap twice so: P(A) + P(B) - P(A and B) → General Addition rule for or

AMS 7
Notes
8/13

Julia Larson

$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \xrightarrow[\text{random (IID)}]{\text{at}} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ $P(Y_1 = 2 \text{ and } Y_2 = 2) = 2$
have to use Random with Replacement (IID)

ELM Works

		Y_2: 1	2	9
Y_1	1	1,1	1,2	1,9
	2	2,1	2,2	2,9
	9	9,1	9,2	9,9

(Joint Distribution)

$P(Y_1 = 2 \text{ and } Y_2 = 2) = \frac{1}{9}$
$P(Y_1 = 2) = \frac{1}{3}$
$P(Y_2 = 2) \ \frac{1}{3}$
$P(Y_1 = 2 \text{ and } Y_2 = 2) = \frac{1}{9} = P(Y_1 = 2) \cdot P(Y_2 = 2)$

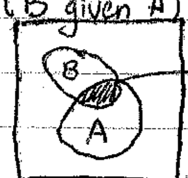
Lets try it with SRS

		Y_2: 1	2	9
Y_1	1	~~1, 1~~	1, 2	1, 9
	2	2, 1	~~2, 2~~	2, 9
	9	9, 1	9, 2	~~9, 9~~

No duplicates

$P(Y_1 = 2 \text{ and } Y_2 = 2) = \frac{0}{6} = 0 \neq \frac{1}{3} \cdot \frac{1}{3}$ does not work with SRS
$P(Y_1 = 2) = \frac{2}{6} = \frac{1}{3}$, $P(Y_2 = 2) = \frac{2}{6} = \frac{1}{3}$

Conditional Probability: Bayes 1720
- $P(B \text{ given } A) = ?$



→ (A and B)

def. $P(B \text{ given } A) = P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$

Multiply by $P(A)$ to get:
$P(A \text{ and } B) = P(A) \cdot P(B|A)$ = General Product Rule for working with (and)
$= P(B) \cdot P(A|B)$

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

Conclusion: IID: 1^{st} draw does not help you to perdict the second draw. def: A, B are independent of each other, 2^{nd} draw behaves as 1^{st}, so $P(A \text{ and } B) = P(A) \cdot P(B)$
SRS: 1^{st} helps predict 2^{nd} (A and B are dependent if knowledge of 1 does not help predict the other [redacted]

def: A and B are independent if: $P(B) = P(B|A)$ and $P(A) = P(A|B)$
proof: $P(A \text{ and } B) = P(A) \cdot P(B|A)$, $\cancel{P(A)} \cdot P(B) = \cancel{P(A)} \cdot P(B|A)$, $P(B) = P(B|A)$