8/23

Homeworks: 30%

Midterm
Final
} 70% (30% + 40%)

Last time:

How to compute # of tigers

Capture/Recapture method

↓

Estimate the size of a wildlife population.

↓

Step 1: 10 tigers are captured, tagged and released.

Step 2: On a later occasion, 20 tigers are captured, and it is found 4 of them are tagged.

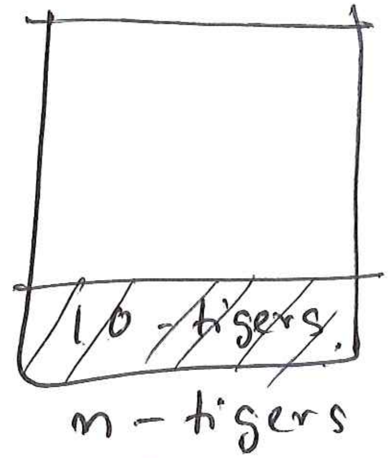
Question: How large is the tiger population?

"Solution"

- We assume that there are
  n tigers.

~~If 10 tigers~~

- 10 of them are tagged. (STEP 1)



- If 20 tigers are captured (in a later occasion). (STEP 2)

This can be done in $\binom{n}{20}$

- Out of these 4 are tagged: (Out of 20 tigers)

$$\binom{10}{4}\binom{n-10}{16}$$

Probability (Out of 20 tigers 4 are tagged) L_n $= \dfrac{\binom{10}{4}\binom{n-10}{16}}{\binom{n}{20}}$

The way n is "estimated"

$\Rightarrow$ Maximize this probability.

(Maximimum likelihood method).

General case:

Out of population (n)

(m (=20)) samples are selected

Out of which (r (=4)) are tagged

[There are a total (t = 10) tagged objects] (STEP 1)

Probability

$$L_n = \frac{\binom{t}{r}\binom{n-t}{m-r}}{\binom{n}{m}}$$

Goal: Maximize L_n

To do this:

$$\frac{L_n}{L_{n-1}} = \frac{\cancel{\binom{t}{r}}\binom{n-t}{m-r}}{\binom{n}{m}} \cdot \frac{\binom{n-1}{m}}{\cancel{\binom{t}{r}}\binom{n-1-t}{m-r}}$$

$$\frac{L_n}{L_{n-1}} = \frac{(n-t)(n-m)}{n(n-t-m+r)}$$ (Check it!)

This ratio > 1 ($\Rightarrow L_n$ is increasing)

$$\frac{(n-t)(n-m)}{n(n-t-m+r)} > 1$$

$$\Rightarrow n^2 - nm - nt + mt > n^2 - nt - nm - nr$$

$$\Rightarrow mt > nr$$

$$\Rightarrow \boxed{\frac{mt}{r} > n}$$

So when $n < \frac{mt}{r}$, L_n is increasing

AND

$\Rightarrow$ If $n > \frac{mt}{r}$, L_n is decreasing.

Maximum of L_n is attained @ $\boxed{n = \frac{mt}{r}}$ (nearest integer)

$$\boxed{\text{The population size} = \frac{mt}{r}}$$

To our last problem:

What is the population size of tiger?

$$= \frac{mt}{r} = \frac{20 * 10}{4} = \boxed{50}$$

[Aside: $\frac{4}{20} = \frac{10}{n} \Rightarrow n = 50$]

Conditional Probability

T+ = Rapid test shows COVID +

T− = Rapid test shows COVID −

D+ = COVID present

D− = COVID absent

	D+	D-	Total.
T+	25	14	39
T-	18	78	96
Total	43	92	135

Table 1

	D+	D-	Total
T+	$\frac{25}{135}=0.185$	$\frac{14}{135}=0.104$	$\frac{39}{135}=0.289$
T-	$\frac{18}{135}=0.133$	$\frac{78}{135}=0.578$	$\frac{96}{135}=0.711$
Total	$\frac{43}{135}=0.318$	$\frac{92}{135}=0.682$	$\frac{135}{135}=1$

Table 2

Total Probability

From the table: (Table 2)

$$\left.\begin{aligned} P(T+) &= 0.289 \\ P(T-) &= 0.711 \\ P(D+) &= 0.318 \\ P(D-) &= 0.682 \end{aligned}\right\} \text{Marginal probabilities}$$

Suppose a test is conducted and we get a positive result (T+) ‖ KNOW!

What is the probability that there is actually ~~the~~ COVID? (D+)

$$P(D+ \mid T+) = \frac{25}{39} = 0.640$$

"GIVEN"

$$= \frac{25/135}{39/135} = \frac{P(D+ \cap T+)}{P(T+)}$$

$$P(D- \mid T-) = \frac{78}{96} = 0.848$$

"GIVEN"

$$= \frac{P(D- \cap T-)}{P(T-)}$$

In general

Conditional prob.

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}}, \text{ if } P(B) \neq 0$$

Multiplication Law

Let A and B be events and assume $P(B) \neq 0$.

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

Example: A box contains

3 red balls
1 blue ball.

Two balls ~~are~~ are to be selected w/o replacement.

What is the prob. that they are both red?

Solution: R_1: The first ball is red.

R_2: The second ball is red.

$$P(R_1 \cap R_2) = P(R_1) \cdot P(R_2 | R_1)$$

$$= \frac{3}{4} * \frac{2}{3} = \frac{1}{2} = \boxed{0.5}$$

[**Aside:**

$$P(R_1 \cap R_2) = P(R_1)\,P(R_2 \mid R_1)$$
$$\|$$
$$\cancel{P(R_2)\cdot P(R_1 \mid R_2)}$$
]

Law of total probability

Let $B_1, B_2, \ldots, B_n$ be such that

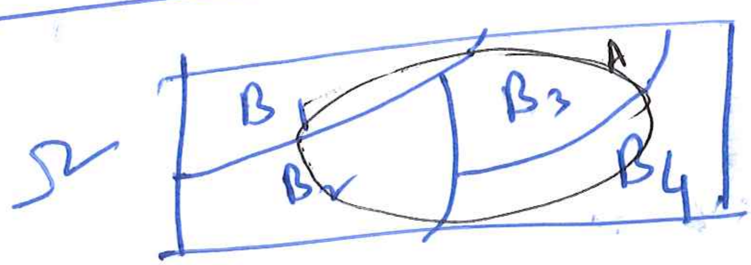
$$\bigcup_{i=1}^{n} B_i = \Omega \leftarrow \text{(sample space)}$$

and $B_i \cap B_j = \phi$, for $i \neq j$

(with $P(B_i) > 0$, for all i)

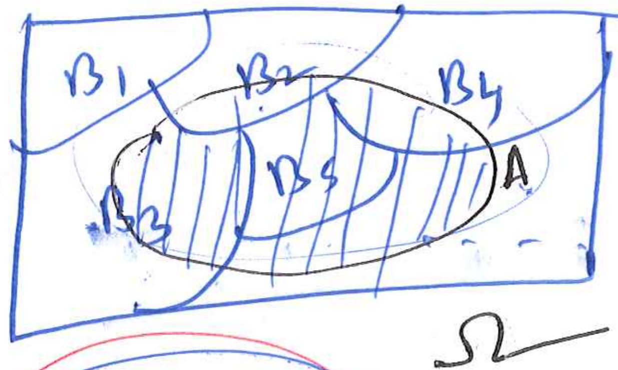
Then for any event A,

$$\boxed{P(A) = \sum_{i=1}^{n} P(A \mid B_i)\, P(B_i)}$$



Proof:

$$P(A) = P(A \cap \Omega)$$



$$= P\left(A \cap \left(\bigcup_{i=1}^{n} B_i\right)\right)$$

$$\Omega = \bigcup_{i=1}^{n} B_i$$

$$= P\left(\bigcup_{i=1}^{n} (A \cap B_i)\right)$$

Now $A \cap B_i$ are disjoint for $i = 1, 2, \dots, n$

By ~~assum~~ axiom (3) of prob

$$= \sum_{i=1}^{n} P(A \cap B_i)$$

$$= \sum_{i=1}^{n} P(A \mid B_i)\, P(B_i)$$

(Done)

Example: Box → 3 red balls, 1 blue ball

(Draw w/o replacement)

What is the prob. that a red ball is selected in the second draw.

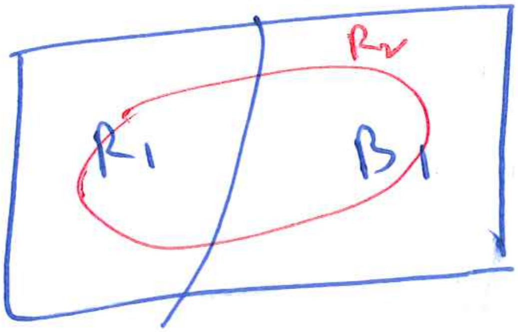
Solution:

$$P(R_2) = P(R_2 | R_1)P(R_1) + P(R_2 | B_1)P(B_1)$$

Law of total prob

$$= \frac{2}{3} * \frac{3}{4} + 1 * \frac{1}{4}$$

$$= \boxed{\frac{3}{4}}$$



Example: Suppose that occupations are grouped into

- U (Upper)
- M (middle)
- L (lower)

U_1 : father's occupation is U

M_1 : . . . . is M

L_1 : . . . . is L

U_2: Son's occupation is U

M_2: - - - - - - - - is M

L_2: - - - - - - - - is L

	U_2	M_2	L_2
U_1	$P(U_2 \mid U_1)$	$P(M_2 \mid U_1)$	$P(L_2 \mid U_1)$
M_1	$P(U_2 \mid M_1)$	$P(M_2 \mid M_1)$	$P(L_2 \mid M_1)$
L_1	$P(U_2 \mid L_1)$	$P(M_2 \mid L_1)$	$P(L_2 \mid L_1)$

Table 1 (Notation)

Transition probability

Prob. of occupation
father → son

Transition matrix

	U_2	M_2	L_2
U_1	0.45	0.48	0.07
M_1	0.05	0.70	0.25
L_1	0.01	0.50	0.49

Table with number (Notation in the LAST table)

[Note the #s are NOT $P(U_1 \cap U_2)$

~~$P(U_1 \cap U_2)$~~ $P(U_2 \mid U_1) = 0.45$.]

It is known

$$P(U_1) = 0.1, \quad P(M_1) = 0.4$$

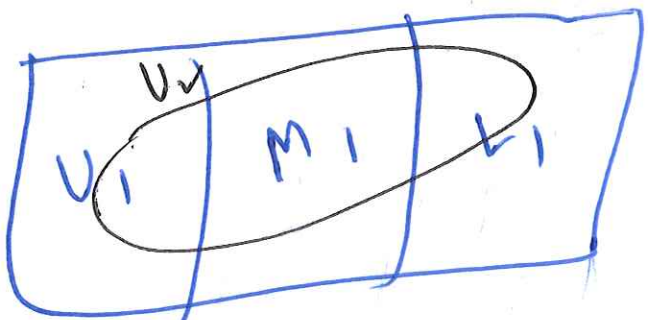
$$P(L_1) = 0.5.$$

Question: $P(U_2)$?

Law of total prob.

Answer:

$$P(U_2) = P(U_2 | U_1) P(U_1) + P(U_2 | M_1) P(M_1) + P(U_2 | L_1) P(L_1)$$



from table

from table.

$$= 0.45 * 0.1 + 0.05 * 0.4 + 0.01 * 0.5$$

$$= 0.07 \; //$$

Think: $P(U_1 | U_2) = \,?$

→ NEXT class!