INTRODUCTION IN PROBABILITIES

Fundamentals

PROBABILITY THEORY

- Probability
  - A numerical measure of uncertainty
  - Quantify the terms “unlikely”, “possible”, “likely”, “probable”, etc.

DEFINITION

- The experiment: putting in practice the condition for studying different situations. Each time when the experiment is done, the result is random

DEFINITION

- The outcome is the result of one execution of the experiment
- The event- the result of an outcome
- The elementary event – the result of JUST one outcome

THE CLASSICAL DEFINITION OF PROBABILITY

- Assuming that for an experiment, all the events are equally plausible, the probability for a particular event is equal with the ratio between the number of the possible ways for that particular event to happen and the total number of outcomes that the experiment could generate

$$P(A) = \frac{s}{n}$$

EXEMPLE

- What is the probability that the head will fall by throwing a coin?
- Possible outcomes: n=2 – (head, tail);
- favorable cases: s= 1 (head);
- probability = 1/2

PROBLEM:

- If we random select a person from a community, what is the chance to have the blood type 0

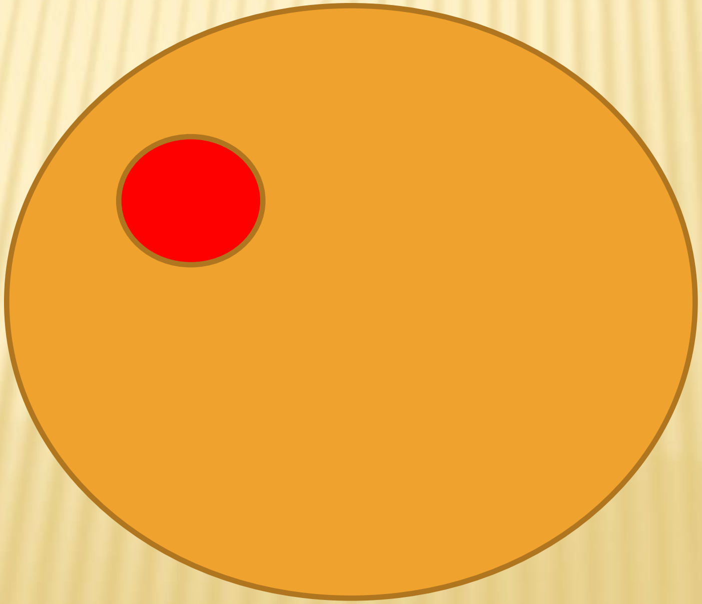
Blood type	**Probability**
0	**0,42**
A	**0,43**
B	**0,11**
AB	**0,04**

THE AXIOMATIC DEFINITION FOR PROBABILITY

- Having an experiment (H) for which E is the set of all possible results. E is called <u>event space</u>.
- Any subset A of E is called the <u>event</u>. If A contains just one element from E, this event will be called <u>an elementary event</u>
- Any event whose realization depends on at least two events is a compound event

EVENTS - SETS

- Event space– the set of all sub-sets
- Events – sub-sets



THE AXIOMATIC DEFINITION FOR PROBABILITY

- The null set ø and the event space E are also events, the impossible event and the certain event

COMPOSED EVENTS (OPERATION WITH EVENTS)

- Union A U B
- Intersection $A \cap B$
- Oposite of A non A
- Incompatble events
- Compatible events
- Involving events

THE AXIOMATIC DEFINITION FOR PROBABILITY

- Having an event space (E) associated to an experiment H and Ω the set of all events (all parts of E).
- The function $P:\Omega \longrightarrow R$ is called probability function, and P(A) is called probability of the event A, if the following axioms are satisfied:
  - $0 \leq P(A) \leq 1$, for any event A of Ω
  - $P(E)=1$
  - If A şi B are incompatible, then

$$P(AUB) = P(A)+P(B)$$

NOTATIONS

- For events : {……}
  - Ex. A={the age of a person is no more than 20 years}
- Probability is a number
  - P(A)=x

PROPERTIES

- If $A_1, A_2, \ldots, A_n$ are incompatible events then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

- P(ø)=0
- P(non A) = 1-P(A)

PROPERTIES

- For any events A and B

$$P(AUB) = P(A) + P(B) - P(A \cap B)$$

ANSWER THE QUESTIONS!

https://forms.office.com/r/29y73XagJF

SITUATION

- Part of the job of a physician is to establish diagnoses. This process starts from the first meeting between the doctor and the patient…

Or before?

SITUATION

- 34 years old female
- Send to a specialty control because of headaches

- 1/3 of those - migraines
- 1/3 – abuse of medication
- 1/3 – anything else – Hypertension, problems of the cervical joints, tumors, etc

SITUATION

- Doing patient interview, clinical examinations, lab tests, the doctor establish the “most likely” diagnose and for it the appropriate treatment
- In time, different factors (events) could intervene, the initial diagnose could be confirmed or another “most likely” diagnose could appear.

INDEPENDENT EVENTS

- Two events A and B are independent if and only if

$$P(A \cap B) = P(\mathrm{A}) \mathrm{x} \mathrm{P}(\mathrm{B})$$

DEPENDENT EVENTS

- Two events A and B are dependent if

$$P(A \cap B) \neq P(\mathrm{A})\mathrm{x}\mathrm{P}(\mathrm{B})$$

APLICATIONS

- In a study about the familial aggregation of hypertension researchers compute the probability of hypertension on men P(A)=0,2 and on women P(B)=0,1
- Compute the probability to have hypertension in the family (at least one person to have hypertension)?

Probability that at least one person in family to have hypertension is the probability that:

- Man has hypertension
- Woman has hypertension
- Both have hypertension

- $P(A \cup B) = P(A)+P(B)-P(A \cap B)=0.2+0.1-0.2*0.1=0.28$

APLICATIONS

- In a study about the familial aggregation of hypertension researchers compute the probability of hypertension on mother P(A)=0,1, on the first child P(B)=0,2 and the frequency of hypertension on children having the mother affected by hypertension $Pr(A \cap B) = 0,5$
- Are those events independent?

If the events are independent then
$P(A \cap B)=P(A)xP(B)$

$P(A)xP(B)=0,1*0,2=0,02$
$P(A \cap B)=0,5$, so the events are not independent

CONDITIONAL PROBABILITIES

- If A and B are two arbitrary events, the conditional probability of A given B- P(A|B) - is the probability of the event A to occur when event B already have occurred.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- If A and B are independent events then $P(B|A)=P(B)$

PROBABILITY OF INTERSECTION

- For any events A and B

$$P(A \cap B) = P(\mathrm{A})\mathrm{x}\mathrm{P}(\mathrm{B}|\mathrm{A}) = \mathrm{P}(\mathrm{B})\mathrm{x}\mathrm{P}(\mathrm{A}|\mathrm{B})$$

BAYES FORMULA

- Considering A and B two NOT independent events. Then

$$P(A|B) = \frac{P(B|A) x P(A)}{P(B)}$$

$$Pr(A\,|\,B) = \frac{Pr(B\,|\,A)\cdot Pr(A)}{Pr(B\,|\,A)\cdot Pr(A) + Pr(B\,|\,nonA)\cdot Pr(nonA)}$$

ANSWER THE QUESTIONS!

https://forms.office.com/r/35Tckt3K4U

MEDICAL APPLICATION OF PROBABILITIES

PROBABILITY

- The relative frequency in the sample is a probability function!

- Age: 19, 19, 20, 21, 20, 19, 19, 20, 21, 20
- F(19)=4/10; F(20)=4/10; F(21)=2/10

PREVALENCE

- If we have a sample of size n in which we study the presence of a disease, this disease being encountered in k people, the prevalence of the disease in the sample is k / n

DEFINITIONS

- Considering T diagnostic test and o D a disease

- A fals positive is a person for which the test is positive but is NOT affected by the disease
- A fals negative is a person for which the test is negative but is affected by the disease

CONTINGENCY TABLE

Disease / Test	D sick	non(D) "healthy"	Total
T Positive test	a (RP)	b (FP)	a+b
non (T) Negative test	c (FN)	d (RN)	c+d
Total	a+c	b+d	n

SENSITIVITY

- **Sensitivity** (also called the *true positive rate*, or the recall rate in some fields) measures the proportion of actual positives which are correctly identified as such (e.g. the percentage of sick people who are correctly identified as having the condition).

$$Se = \Pr(T/D) = \frac{RP}{RP+FN} = \frac{a}{a+c} = \frac{\frac{a}{n}}{\frac{a+c}{n}} = \frac{\Pr(T \cap D)}{\Pr(D)}$$

Disease / Test	D sick	non(D) “healthy”	Total
T Positive test	a (RP)	b (FP)	a+b
non (T) Negative test	c (FN)	d (RN)	c+d
Disease / Test	a+c	b+d	n

SPECIFICITY

- **Specificity** measures the proportion of negatives which are correctly identified as such (e.g. the percentage of healthy people who are correctly identified as not having the condition, sometimes called the *true negative rate*).

$$Sp = \Pr(non(T)/non(D)) = \frac{RN}{FP+RN} = \frac{d}{b+d} = \frac{\frac{d}{n}}{\frac{b+d}{n}} = \frac{\Pr(nonT \cap nonD)}{\Pr(nonD)}$$

Disease / Test	D sick	non(D) "healthy"	Total
T Positive test	a (RP)	b (FP)	a+b
non (T) Negative test	c (FN)	d (RN)	c+d
Disease / Test	a+c	b+d	n

POSITIVE PREDICTIVE VALUE

- the positive predictive value, or precision rate is the proportion of positive test results that are true positives (such as correct diagnoses)

$$VPP = \Pr(D/T) = \frac{\Pr(T \cap D)}{\Pr(D)} = \frac{RP}{RP+FP} = \frac{a}{a+b}$$

Disease / Test	D sick	non(D) "healthy"	Total
T Positive test	a (RP)	b (FP)	a+b
non (T) Negative test	c (FN)	d (RN)	c+d
Disease / Test	a+c	b+d	n

NEGATIVE PREDICTIVE VALUE

- proportion of subjects with a negative test result who are correctly diagnosed

$$VPN = \Pr(nonD / nonT) = \frac{\Pr(nonD \cap nonT)}{\Pr(nonT)} = \frac{RN}{FN + RN} = \frac{d}{c+d}$$

Disease / Test	D sick	non(D) “healthy”	Total
T Positive test	a (RP)	b (FP)	a+b
non (T) Negative test	c (FN)	d (RN)	c+d
Disease / Test	a+c	b+d	n

RELATION BETWEEN SE, SP, PREVALENCE (PR), PPV, NPV

$$PPV = \frac{Se * \; Pr}{Se * \mathrm{Pr} + (1 - Sp) * (1 - \mathrm{Pr})}$$

$$NPV = \frac{Sp * (1 - \mathrm{Pr})}{Sp * (1 - \mathrm{Pr}) + (1 - Se) * Pr}$$

APLICATION

- Automated systems for the determination of hypertensive disease have been introduced in pharmacies.This system indicates 84% of hypertensive persons and 23% of normotensive persons having hypertensive disease
- If 20% of the population ìs hypertensive, compute PPV and NPV for this system?
- Se = 0,84
- Sp = 1 – 0,23 = 0,77

- PPV=0,84x0,2/(0,84x0,2+0,23x0,8)=0,168/0,352=0,48
- NPV=0,77x0,8/(0,77x0,8+0,16*0,2)=0,616/0,648=0,95

ANSWER THE QUESTIONS!

https://forms.office.com/r/YT34WqQEZz

THANK YOU!

- Please evaluate this lecture!

https://forms.office.com/r/6QFxLD4GD4