Topic 4: Probability Theory

Reference: Chapter 4

Objectives:

(i) **Alternative definitions of ___________**

►O________ (a priori) Probability

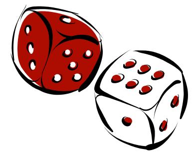
►Long-Run Relative Frequency

►__________ Probability

(ii) **Counting rules** – Permutations

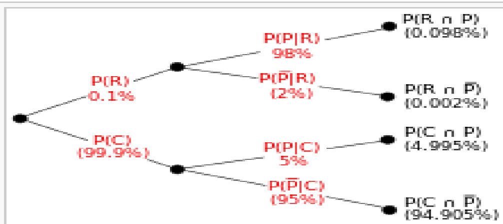
-- Combinations

♥♦♣♠ ♥♦♠♣ ♥♣♦♠



(iii) **Probability rules**

(iv) Demonstrate how probability is used to assist in measuring “uncertain” behaviour.



Tree diagram illustrating frequentist example. R, C, P and P bar are the events representing rare, common, pattern and no pattern. Percentages in parentheses are calculated. Note that three independent values are given, so it is possible to calculate the inverse tree (see figure above).

Consider all possible outcomes of some uncertain situation:

⇨**Event:** an event is some subset of all the possible outcomes in a decision-making situation under conditions of ***un*__________**.

Example: Toss $2 coin –head / tail

⇨Do this twice

Outcomes: {_,_},{_,_}{H,T}, {T,H}.

Event: ***"Exactly one tail".***

⇨{T,H}{H,T} ⇦(2 out of 4 outcomes)

⇨ **<u>Probability</u>**: is a number between **z____** and **o__** that indicates how likely it is that an event will occur.

- If an event is **impossible**, its probability is <u>zero</u>.
- If an event is **certain**, its probability is <u>one</u>.

A fundamental part of statistical analysis includes finding the value that represents how likely an event will occur.

⇨ **Probability theory provides this.**

<u>Example:</u> Toss a coin:

<u>Event 1:</u>
"At least one head, or one tail."
(Certain)

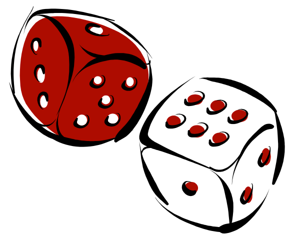
$P(E_1) = 1;$

$P(E_2) = 0.$

<u>Event 2:</u>
"No heads and no tails."
(Impossible)

● An **<u>Experiment</u>**: is a situation we can <u>replicate</u> under essentially _____ conditions.

<u>Example</u>: Rolling pair of dice:



Different replications of the experiment may result in different outcomes.

- Each outcome is a "**sample point**."
- Set of all sample points is the "**sample space.**"

❖Must be careful to ensure that outcomes are mutually exclusive and ___________.

⇨**<u>Mutually exclusive</u>** outcomes have no ________.

⇨**<u>Exhaustive outcomes</u>** mean that no possible outcome is ____ off the list of outcomes.

<u>Example:</u> Toss two $2 coins: 

Sample space is: (H,H) (T,T), (H,T) , (T,H).

⇨All possible events have been accounted.

****

There is some disagreement about the definitions of the probability of an event.

❖Several ways to define probability:

(A) Objective (a priori) Probability:

$$\text{Probability of Event} = \left\{ \frac{\#\text{ of outcomes in favour}}{\#\text{ of outcomes in total}} \right\}$$

❖Determined by objective _________ and would have the same value **regardles**s of who did the interpretation.

Eg. P(one head) $= \frac{2}{4} = \frac{1}{2}$.

(B) Long-Run ________ Frequency:

► Replicate experiment many times.

► Keep track of the proportion of favourable outcomes – look at the **limiting** _____.

Example: Toss two coins. The probability of **exactly** one tail:

(H,H) (T,T), (H,T), (T,H), (H,H) (T,T), (H,T), (T,H),(H,H), ...

$$\text{Prob} = \frac{0}{1}, \frac{0}{2}, \frac{1}{3}, \frac{2}{4}, \frac{2}{5}, \frac{2}{6}, \cdots$$

STEPS Relative Frequency

Relative frequency is another term for proportion; it is the value calculated by dividing the number of times an event occurs by the total number of times an experiment is carried out. The probability of an event can be thought of as its long-run relative frequency when the experiment is carried out many times.

If an experiment is repeated n times, and event E occurs r times, then the relative frequency of the event E is defined to be

$rf_n(E) = r/n$

Example

Experiment: Tossing a fair coin 50 times (n = 50)

Event E = 'heads'

Result: 30 heads, 20 tails, so r = 30

Relative frequency: $rf_n(E) = r/n = 30/50 = 3/5 = 0.6$

If an experiment is repeated many, many times without changing the experimental conditions, the relative frequency of any particular event will settle down to some value. The probability of the event can be defined as the limiting value of the relative frequency:

$P(E) = \lim_{n \to \infty} rf_n(E)$

For example, in the above experiment, the relative frequency of the event 'heads' will settle down to a value of approximately 0.5 if the experiment is repeated many more times.

Thus, if n_t is the total number of trials and n_x is the number of trials where the event x occurred, the probability $P(x)$ of the event occurring will be approximated by the relative frequency as follows:

$$P(x) \approx \frac{n_x}{n_t}.$$

A further and more controversial claim is that in the "long run," as the number of trials approaches infinity, the relative frequency will converge exactly to the probability:[1]

$$P(x) = \lim_{n_t \to \infty} \frac{n_x}{n_t}.$$

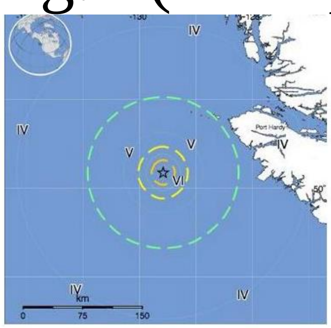
(C) Subjective Probability:

❖Assigns probabilities based on the ________-maker's **subjective** estimates, using _____ knowledge, information, and experience as a guide.

❖Subjective (personal) "degree of belief."

❖Useful for "once-and-for-all" events.

Eg. P(earthquake during the next exam)



Update: New quakes today off Vancouver Island

Five more earthquakes registered off B.C. coast.

However probability is defined, it satisfies:

$$(i)\quad 0 \le P(E_i) \le 1; \qquad All\ E_i\ in\ sample\ space(S).$$

$$(ii)\ \sum P(E_i) = P(S) = P(E_1) + P(E_2) + \ldots = 1$$

All events are ________ exclusive and exhaustive.

Christiaan Huygens published the first book on probability

Carl Friedrich Gauss

Independent probability [edit]

If two events, *A* and *B* are independent then the joint probability is

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

for example, if two coins are flipped the chance of both being heads is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.[19]

Mutually exclusive [edit]

If either event *A* or event *B* or both events occur on a single performance of an experiment this is called the union of the events *A* and *B* denoted as $P(A \cup B)$. If two events are mutually exclusive then the probability of either occurring is

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B).$$

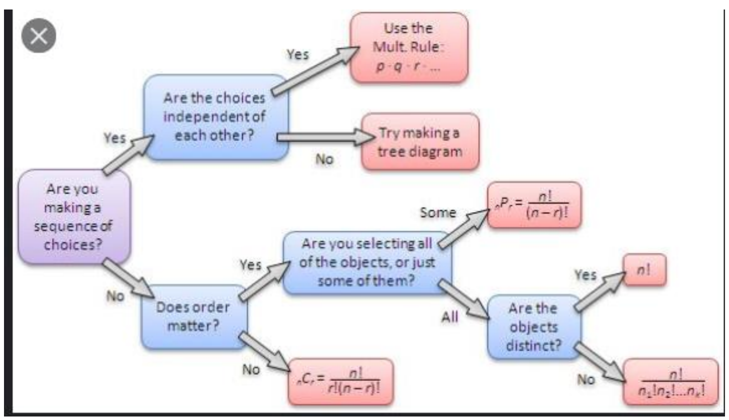
For example, the chance of rolling a 1 or 2 on a six-sided die is $P(1 \text{ or } 2) = P(1) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$

To compute probabilities we need to be able to define the sample space, S, properly. This requires us to determine all outcomes.

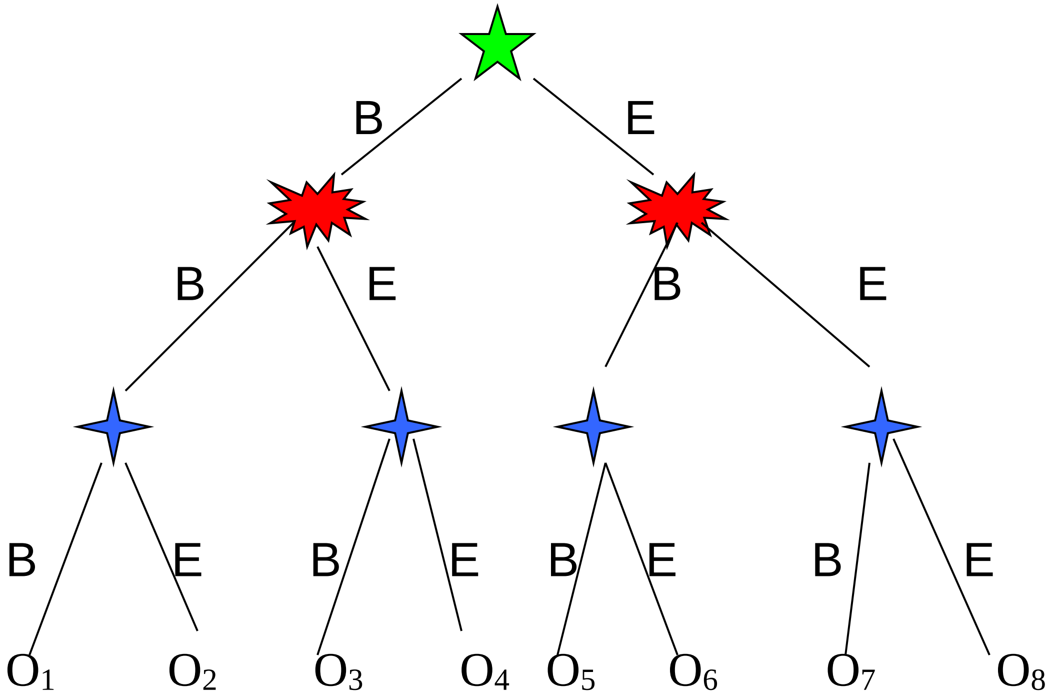
➡Need to understand certain "**counting rules.**"

Counting Rules

●Most experiment involve several steps, each involving several possible outcomes:



Example: Select 3 students: (Business / Economics)



⇨Eight outcomes in this sample space.

[# of outcomes =(2)(2)(2)=8]

❖If business/economics equally likely at each step, then each outcome has probability of __.

Probability of Event= Sum of Probability of related outcomes.

Example: P[2 _________] $= P(O_2 \text{ or } O_3 \text{ or } O_5) = \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) = \frac{3}{8}$

Permutations and Combinations:

In some probability problems where we have a set of distinguishable objects, we may want to know how many different ways there are of __________these objects.

Example:

Brands of squash shoes: Nike (n), Reebok (R) and Addidas (A)

- ⇨Can choose most __________ in 3 ways
- ⇨Can choose ____ in 2 ways
- ⇨Can choose last in 1 way.

So, the number of orderings is $(3\times 2\times 1)=6$

In general, if you have "n" such items, the number of distinct arrangements is:

$$n(n-1)(n-2)\ldots 3 \times 2 \times 1 = n!$$

↳'n f_________'

(Note: 0!=1 by definition)

In the above example, the <u> 6 orderings</u> are:

(N, R, A) (N, A, R) (R, A, N)
(R, N, A) (A, R, N) (A, N, R)

Another name for such orderings is **"p___________."**

⇨(N, R, A) is one of the 6 possible permutations.

- Sometimes we are interested in taking just a **<u>subset</u>** of the total number of items, and considering permutations of them.

For <u>example</u>: 4 brands of shoes – want to choose just 2, and order them:

⇨Can choose 1st in 4 ways.
⇨Can choose 2nd in 3 ways.

So, number of permutations is (4)(3)=12

For all 4, # of permutations=4!=24

Notation:

nPx=(# of permutations of n items taken X at a time).

For the last example: nPx=n(n-1)(n-2) … (n - x + 1)

Note:

$$\frac{n!}{(n-x)!} = \left[\frac{n(n-1)(n-2)\cdots(3)(2)(1)}{(n-2)(n-x-1)\cdots(3)(2)(1)} \right]$$

$$= n(n-1)(n-2)\cdots(n-x+1)$$

$$= nPx$$

<u>Example</u>: 10 job applicants; 2 positions (select 2 people from 10).

Number of Permutations:(10)(9)=__P_.

In evaluating permutations, the <u>order</u> of the items ________.

For example, for 3 brands of squash shoes, the permutations of all 3 were:

(N, R, A) (N, A, R) (R, A, N)
(R, N, A) (A, R, N) (A, N, R)

Also for this problem choosing 2 shoes:

$\# = {}_3P_2 = (3)(2) = 6.$

(N, R) (A, R) (R, A)
(R, N) (N, A) (A, N,)

****

In some situations, the ______ **is not relevant**.

For example, when choosing 2 people from 10 applicants, you just want a team of 2 (no need to rank).

C____________ of objects are groupings in which order is irrelevant.

<u>Notation:</u> **nCx** = (# of combinations of n items, X at a time).

For a fixed n and X, nCx < nPx.

In fact:

$$nCx = \left(\frac{nPx}{x!} \right) = \frac{n!}{(n-x)!x!}$$

<u>Example</u>: Choose 2 people for jobs from 4 applicants.

<u>(A) Order Matters</u>: number of ways =_*P*_=(4)(3)=12

(A, B) (B, A) (A, C) (C, A)

(A, D) (D, A) (B, C) (C, B)

(B, D) (D, B) (C, D) (D, C)

<u>(B) Order Irrelevant:</u> disregard ____:

(A, B) (A, C)

(A, D) (B, C)

(B, D) (C, D)

$$\# \text{ of ways} = {}_4C_2 = \frac{{}_4P_2}{2!} = \frac{12}{2} = 6.$$

More Examples:

Select code for a gadget, by setting 9 switches to "+" or "-".

If you forget the code, how many attempts are needed?

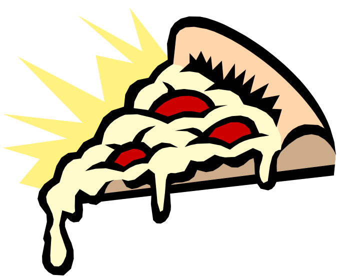
$$\# = (2)(2)\cdots(2) = 2^9 = 512$$

What is the probability of being correct on the first attempt?

$$= \tfrac{1}{512} \Leftarrow \text{ same on any attempt.}$$

Example #2:

Can order pizza with none or up to 8 toppings. How many combinations of toppings are possible? For each topping, either "yes" or "no." (2 possibilities)



Possibilities=(2) (2) …(2)=2^8=____.

Example #3:

The professor gives you __sample questions, 2 of which will be chosen randomly for the mid-term test. You only have time to study 2 questions:

Sample space: (1, 2) (1, 3) (1, 4) (1, 5) (2, 3)
(2, 4)(2, 5) (3 ,4) (3, 5) (4, 5)

(No replacement ; order irrelevant)

$$\# = {}_5C_2 = \frac{5!}{3!2!} = \left[\frac{5 \times 4}{2 \times 1}\right] = 10$$

What is the probability you study the correct 2 questions?

Suppose she has already chosen the 2 questions.

This fixed combination is just one of 10 possibilities.

You have a 10% probability of choosing correctly.

Probability Rules:

We know the probability of an _____ is the sum of the probabilities of outcomes making up that event.

Can often evaluate the probability of an event from probabilities of other events.

I.e. We can often determine the probability of an event from the knowledge about the probability of one or more other events in the sample space.

Recall: An event is a set -- a subset of the sample space.

Example: Toss 2 coins.

S={(H, H) (T, T) (H, T) (T, H)}

E={exactly ____head} ={(T, H) (H, T)}

and clearly $\boxed{E \in S}$.

So ⇨ $\boxed{P(E) = \frac{2}{4} = \frac{1}{2}.}$

If $\overline{E}$ is the **c__________** of E, then:

⇨ $\boxed{P(\overline{E}) = 1 - P(E)}$ because P(S)=1 and $\left(E \cup \overline{E}\right) = S$.

(Note: E and $\overline{E}$ are mutually exclusive and exhaustive events.)

Probability Rules: Basic Definitions:

If we let A and B represent two events of interest in a particular experiment:

(1) $P(\overline{A})$=Probability that A does not occur in one trial of the experiment.

$P(\overline{A})$ = probability of the ___________ of A.

(2) P(A|B)=Prob A occurs **given** that B has taken place.

"____________ probability"

(3) $P(A \cap B)$ = Probability that **both** A and B occur in one trial of the experiment.

Intersection or______ probability of A and B.

(4) $P(A \cup B)$ =Probability that either A or B or both occur in one trial of the experiment.

Probability of the _____ of A and B.

Conditional Probability

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

We want to consider the probability of some event occurring, given that (conditional on) some other event has occurred (will occur).

Let A and B represent events.

P(A|B)=P(Event A given Event B)

Example: Revenue Canada conducts an audit of certain type of firms in 2 provinces. There are __,___ firms in total; 12,000 in Quebec and 8,000 in Alberta. Suppose 2000 Quebec firms and ___ Alberta firms "avoid" tax.

Revenue Canada samples a firm for audit. What is the probability it is a tax avoider, given that it is a Quebec firm?

$$P(A|Q) = \frac{2000}{12{,}000} = \frac{1}{6}$$

$$P(A|Q) = \frac{\left(2000 \big/ 20{,}000\right)}{\left(12{,}000 \big/ 20{,}000\right)}$$

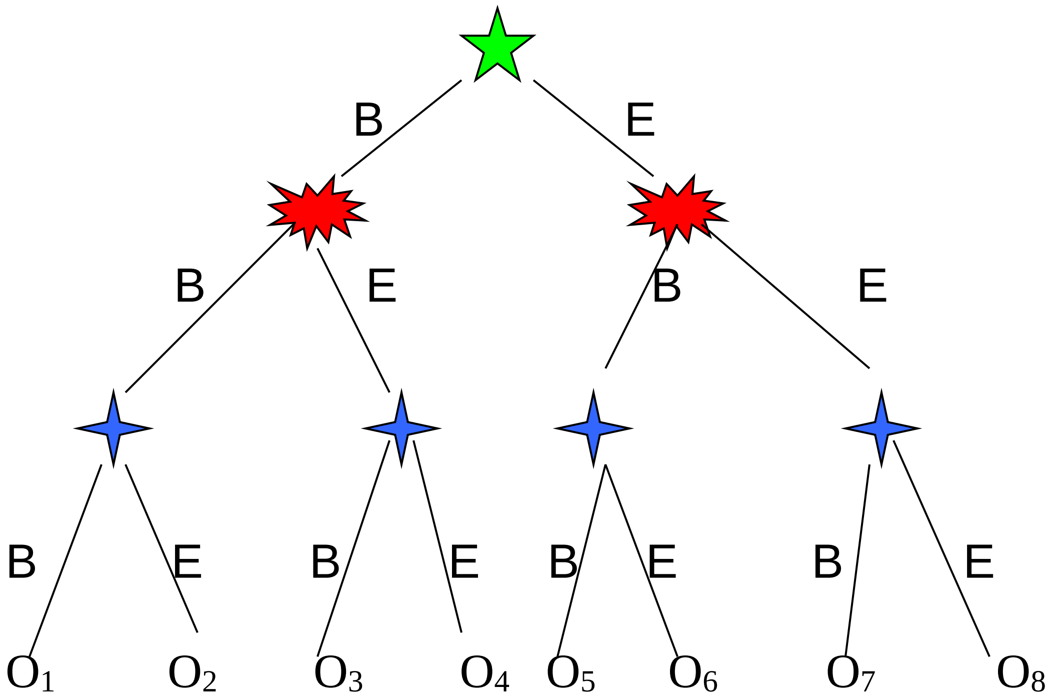
$$= \frac{P(A \text{ and } Q)}{P(Q)} = \frac{P(A \cap Q)}{P(Q)}$$

Where P(Q) is the conditional probability

S=All firms=20,000

Compare: P(tax avoiding firm) = $P(A) = \frac{2500}{20,000} = \frac{1}{8} \neq \frac{1}{6}$.

Example: Business/Economics from before:



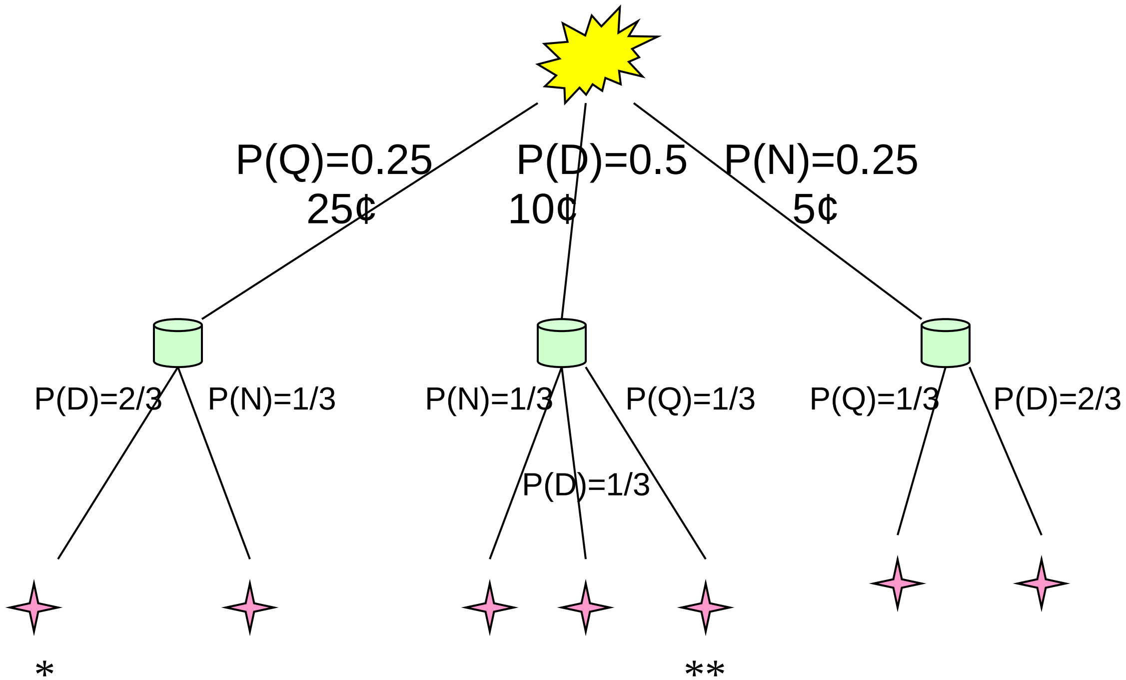
P(choosing 2 business)= $\frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$

P(choosing 2 business | first choice is business) = P[(BBE) or (BEB)]= $\frac{2}{4} = \frac{1}{2}$.

Rule: $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/8}{1/2} = \frac{4}{8} = \frac{1}{2}.$$

$$P(\text{2 Business}|\text{Business First}) = \frac{P(\text{2 Business} \cap \text{Business first})}{P(\text{Business First})} = \frac{2/8}{1/2} = \frac{4}{8} = \frac{1}{2}.$$

Example: Box contains 1 nickel, 2 _____ and 1 _______. Select 2 coins. What is the probability of a dime and quarter? P(Dime and Quarter)?



(Order does not matter.)

P(Q and D)

$$P(Q, D) = P(D|Q) \times P(Q)$$

$$= \left(\frac{2}{3}\right) \times \left(\frac{1}{4}\right) = \frac{2}{12} = \frac{1}{6}.$$

$$P(D, Q) = P(Q|D) \times P(D)$$

$$= \left(\frac{1}{3}\right) \times \left(\frac{1}{2}\right) = \frac{1}{6}.$$

So, P(Quarter and dime)= $P(Q \text{ and } D) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$

$$P(D \cap Q) = P(D|Q)P(Q) + P(Q|D)P(D)$$

Two Special Probabilities

From the definition of conditional probability we have:

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

Re-arrange this expression:

(i) $P(A \cap B) = P(A|B)P(B)$ $P(A \cap B) = P(B \cap A)$

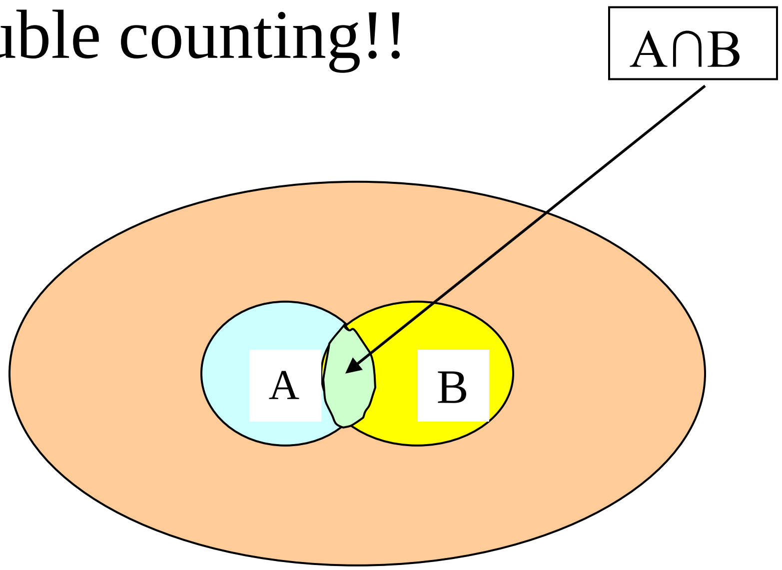
so,

$P(A \cap B)$ also equals $P(B|A)P(A)$.

(ii) Consider $P(A \cup B)$:

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

Avoid double counting!!



Example: Firm Auditing (previous)

P(Avoids tax and in Quebec):

=P(Avoids|Quebec) * P(Quebec)

$$P(A \cap Q) = \left(\frac{1}{6}\right)\left(\frac{12}{20}\right) = \frac{2000}{20{,}000} = \frac{1}{10}.$$

$$\left(\frac{2000}{12{,}000}\right)\left(\frac{12{,}000}{20{,}000}\right)$$

P(Avoids or in Quebec):

=P(Avoids)+P(Quebec)-P(Avoids and Quebec)

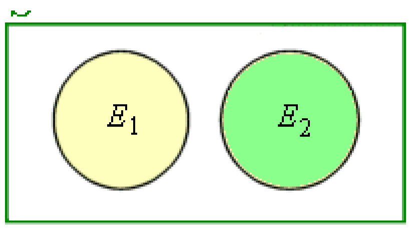
$$P(A \cup Q) = \frac{2500}{20{,}000} + \frac{12{,}000}{20{,}000} - \frac{1}{10} = \frac{12{,}500}{20{,}000} = 0.625$$

Some Special Cases:

(i) Suppose A and B are **mutually exclusive** (so $(A \cap B) = \varnothing$).

Generally: $P(A \cup B) = P(A) + P(B) - (A \cap B)$.

In this case: $P(A \cup B) = P(A) + P(B)$ because $P(A \cap B) = 0$.



Two events are **mutually exclusive** if they cannot occur at the same time.

(ii) **"Independent" Events**:

If P(A) is un_________ by the knowledge of B, and vice versa, then A and B are said to be statistically "independent" events.

If A and B are independent, then:

P(A|B)= ____
P(B|A)=_____

So, in this special case:

$$P(A \cap B) = P(A|B)\,P(B) = P(A) \times P(B).$$

Example: Auditing Firms: To determine whether Avoidance and Quebec are independent event:

P(Avoid|Quebec)= $\frac{1}{6} = 0.1667$

P(Avoid)= $\frac{2500}{20,000} = 0.125$

These 2 events are ___ independent (and not M.E. either).

***Note:* Independence and mutually exclusive are <u>not</u> the same thing.**

(i) Mutually exclusive ⇨ _____________

(ii) Not mutually exclusive: could be independent or dependent.

(iii) Independent ⇨ _______be mutually exclusive.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{P(B)} \neq P(A)$$

Exclusive ⇒ dependent

Mutually exclusive implies dependence, and independence implies not mutually exclusive, but no other simple implications among these conditions holds true.

Example:

A= You are employed by Western Forest Products
B= You are not employed by Western Forest Products
C=The softwood tariff has decreased
D=My shoe size is 17

_ and _ are mutually exclusive and dependent; only one can occur at a time.

A and _ are dependent and not M.E.; both can happen.

A and _ are not M.E. and independent.

Marginal Probabilities:

In many circumstances, it is convenient to assume that a single event always occurs jointly with other events:

Consider the situation:

	B.Sc B.A. B.Ed.	
Male Female	8 20 20	
		100= total #

Convert to Probabilities:

Marginal probability

	B.Sc	**B.A.**	**B.Ed.**	**P(Gender)**
Male				
Female	0.08	0.20	0.20	0.48
P(Degree)	0.18	0.32	0.50	1.00

P(B.A.)=P(B.A. and Male) + P(B.A. and Female)
= 0.12 + 0.20 =0.32

P(Male)= P(Male ∩BSc.)+ P(Male ∩ B.A.) +P(Male ∩ B.Ed.)
= 0.52

Notice that:

$$P(\text{Male} \cap \text{BSc.}) = \underbrace{P(\text{Male} | \text{B.Sc})\, P(\text{B.Sc})},$$

⇨ So, $\underbrace{P(\text{Male}) =}$

$P(\text{Male}|\text{B.Sc})P(\text{B.Sc}) + P(\text{Male}|\text{B.A.})P(\text{B.A.}) + P(\text{Male}|\text{B.Ed.})P(\text{B.Ed.})$

$$P(\text{Male}) = \sum P(\text{Male} \cap \text{Degree})$$

$$= \sum P(\text{Male} | \text{Degree}) P(\text{Degree})$$

$$P(\text{Male}) = P(\text{Male}|\text{B.Sc})P(\text{B.Sc}) + P(\text{Male}|\text{B.A.})P(\text{B.A.}) + P(\text{Male}|\text{B.Ed.})P(\text{B.Ed.})$$

$$= \left[\frac{M \cap BSc}{BSc}\right](BSc) + \left[\frac{M \cap BA}{BA}\right](BA) + \left[\frac{M \cap B.Ed}{B.Ed}\right](B.Ed)$$

$$= \left[\frac{0.10}{0.18}\right](0.18) + \left[\frac{0.12}{0.32}\right](0.32) + \left[\frac{0.30}{0.50}\right](0.50) = 0.52$$

Let A represent an event, such that one of the mutually exclusive events E_1, E_2, …, E_K always must occur jointly with any occurrence of A:

Marginal probability:

$$P(A) = \sum_{i=1}^{K} P(A \cap E_i) = \sum_{i-1}^{K} P(E_i)P(A| E_i)$$

$$= P(E_1)P(A| E_1) + P(E_2)P(A| E_2) + \cdots P(E_K)P(A| E_K)$$

Bayes' Rule

An application of the rules of probability theory involves estimating unknown probabilities and making decisions on the basis of ___ sample information.

Bayesian approach calculates conditional probabilities: $P(E_i|A)$, where A is some new information.

Hence, **Bayes' Rule** is concerned with determining the probability of an event given certain new information: A way of updating probabilities.

$$P(E_i | A) = \frac{P(E_i)P(A|E_i)}{P(A)} = \frac{P(E_i)P(A|E_i)}{\sum_{i=1}^{K} P(A \cap E_i)} = \frac{P(E_i)P(A|E_i)}{\sum_{i=1}^{K} P(E_i)P(A|E_i)}$$

where the E_i's are mutually exclusive and exhaustive.

Question 3.39: Consider two types of economic stabilization policies – fiscal and monetary policy. Assume that the policy decisions made by these two institutions are ____________ of one another and that the action of either group is correct 80% of the time. Finally, assume that the probabilities that the economy follows a generally stable growth pattern due to theses policy actions are:

P(Stable growth | Neither acting correctly = 0.___
P(Stable growth | Both acting correctly) = 0.___
P(Stable growth | Only 1 acting correctly) = 0.___

A) Use the independence assumption to calculate:

(I) P(Neither acting correctly)

$$P(\text{Neither correct}) = P(\text{Mont. NOT Correct}) \times (\text{Fiscal NOT Correct})$$
$$= (0.2) \times (0.2) = 0.04$$

rule: $P(\overline{A} \cap \overline{B}) = P(\overline{A}) \times P(\overline{B})$

(II) P(Both acting correctly)

$$P(\text{Both correct}) = P(\text{Mont. Correct}) \times (\text{Fiscal Correct})$$
$$= (0.8) \times (0.8) = 0.64$$

rule: $P(A \cap B) = P(A) \times P(B)$

(III) P(Only 1 acting correctly)

$$P(\text{One correct}) = P(\text{Mont. NOT Correct}) \times (\text{Fiscal Correct}) + P(\text{Mont. Correct}) \times (\text{Fiscal NOT Correct})$$
$$= (0.2)(0.8) \times (0.8)(0.2) = 0.32$$

rule: $P(\overline{A} \cap B) + P(A \cap \overline{B})$

B) **You are given the sample information that growth is stable for a particular period. Use Bayes' rule to calculate**:

(I) P(Only 1 acting correctly | Stable growth)

$$= \frac{P(\text{Growth}|\text{One correct}) \times P(\text{One correct})}{P(\text{Growth})}$$

$$= \frac{P(\text{Growth}|\text{One correct}) \times P(\text{One correct})}{P(G|0)P(0) + P(G|1)P(1) + P(G|2)P(2)}$$

$$= \frac{(0.7)(0.32)}{(0.4)(0.04) + (0.7)(0.32) + (0.99)(0.64)} = 0.2564$$

(II) P(Both acting correctly| Stable growth)

$$= \frac{P(\text{Growth}|\text{Both correct}) \times P(\text{Both correct})}{P(\text{Growth})}$$

$$= \frac{P(\text{Growth}|\text{Both correct}) \times P(\text{Both correct})}{P(G|0)P(0) + P(G|1)P(1) + P(G|2)P(2)}$$

$$= \frac{(0.99)(0.64)}{(0.4)(0.04) + (0.7)(0.32) + (0.99)(0.64)} = 0.7253$$

(III) P(Neither acting correctly| Stable growth)

$$= \frac{P(\text{Growth}|\text{Neither correct}) \times P(\text{Neither correct})}{P(\text{Growth})}$$

$$= \frac{P(\text{Growth}|\text{Neither correct}) \times P(\text{Neither correct})}{P(G|0)P(0) + P(G|1)P(1) + P(G|2)P(2)}$$

$$= \frac{(0.4)(0.04)}{(0.4)(0.04) + (0.7)(0.32) + (0.99)(0.64)} = 0.0183$$

For general events

Bayes' theorem may be derived from the definition of conditional probability:

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

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

$$\Longrightarrow \; P(A \cap B) = P(A|B)\,P(B) = P(B|A)\,P(A).$$

$$\Longrightarrow \; P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

Simple example

An entomologist spots what might be a rare subspecies of beetle, due to the pattern on its back. In the rare subspecies, 98% have the pattern. In the common subspecies, 5% have a similar pattern, but he cannot distinguish these from memory. The rare subspecies accounts for only 0.1% of the population. How likely is the beetle to be rare?

From the extended form of Bayes' theorem,

$$P(\text{Rare}|\text{Pattern}) = \frac{P(\text{Pattern}|\text{Rare})P(\text{Rare})}{P(\text{Pattern}|\text{Rare})P(\text{Rare}) + P(\text{Pattern}|\text{Common})P(\text{Common})}$$

$$= \frac{0.98 \times 0.001}{0.98 \times 0.001 + 0.05 \times 0.999}$$

$$= 1.9\%$$

Bayes' theorem (also known as Bayes' rule) is a useful tool for calculating conditional probabilities. Bayes' theorem can be stated as follows:

Bayes' theorem. Let $A_1, A_2, \ldots, A_n$ be a set of mutually exclusive events that together form the sample space S. Let B be any event from the same sample space, such that P(B) > 0. Then,

$$P(A_k \mid B) = \frac{P(A_k \cap B)}{P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_n \cap B)}$$

Note: Invoking the fact that $P(A_k \cap B) = P(A_k)P(B \mid A_k)$, Baye's theorem can also be expressed as

$$P(A_k \mid B) = \frac{P(A_k)\,P(B \mid A_k)}{P(A_1)\,P(B \mid A_1) + P(A_2)\,P(B \mid A_2) + \ldots + P(A_n)\,P(B \mid A_n)}$$

Unless you are a world-class statiscian, Bayes' theorem (as expressed above) can be intimidating.