Basics of Probability

* spelling errors may occur because no squiggly lines come up.

1) Probability

(a) probability = chance = likelihood = possibility (P1)

(b) probability = chance that something will occur in the unknown future

or

Numerical measure of the likelihood that an event will occur

(c) Let:

E_i = i^{th} experimental outcome (sample point)

$P(E_i)$ = probability of Event i (like roll 4 on die $P(4) = \frac{1}{6}$)

(d) $$0 \le P(E_i) \le 1$$

(e) $P(E_i)$ is never known with certainty (only estimate)

$P(E_i)$ is an estimate of an event that <u>may</u> occur in future

(f) Examples of $P(E_i)$

$P(\text{Roll 4 with die}) = \frac{1}{6}$  6 sides, one "4"

$P(\text{it will Rain in Seattle next year}) \approx 1$

$P(\text{Sun will go out in 1 minute}) \approx 0$

$P(\text{getting 3.5 Grade}) = \frac{2}{35}$ (Based on past Data)

2 Random Experiment

a

1. A process that generates well-defined outcomes.
2. On any single repetition or trial of an experiment, one and only one of the possible experimental outcomes can occur.
3. The experimental outcome that occurs on any trial is determined solely by chance.

b Examples:

Experiment	Experimental outcomes (sample points)
1. Toss a coin ⇒	Head, Tail
2. Drive on Bridge ⇒	Stuck in Traffic, Not stuck
3. Roll Die ⇒	1, 2, 3, 4, 5, 6
4. Make Product ⇒	Defect, Not Defect

3 Multi-step Random Experiment

a. Experiment with more than 1 step

b. Examples:

Not multi-step Experiment: Flip coin 1 time

Multi-step Random Experiment:

1. Roll die 2 times
2. Flip coin 3 times
3. Drive across bridge 7 times

4) Sample point = Experimental Outcome

a) One of the possible experimental outcomes

b) Examples:

1) Experiment: Flip coin 1 time

Sample Point 1 = Head
Sample Point 2 = Tail

2) Experiment: Flip coin 2 times

Sample Point 1 = Head, Head
Sample Point 2 = Head, Tail
Sample Point 3 = Tail, Head
Sample Point 4 = Tail, Tail

c) Sample point is one element of sample space

d) "Sample point" and "experimental outcome" are synonyms.

5) Sample space = S

a) List of all possible experimental outcomes (sample points) for a random experiment

b) Examples:

$$\left\{\begin{matrix}\text{sample space for} \\ \text{Flip coin 1 time}\end{matrix}\right\} = S = \{H, T\}$$

$$\left\{\begin{matrix}\text{sample space for} \\ \text{Flip coin 2 times}\end{matrix}\right\} = S = \{(H,H), (H,T), (T,H), (T,T)\}$$

c) It is not always possible to list all sample points

6 Counting Rule For Multi-step Random Experiment

(a) $\begin{Bmatrix}\text{total Number of}\\ \text{Experimental outcomes}\\ \text{(sample points)}\end{Bmatrix} = n_1 * n_2 * \dots n_k = \begin{Bmatrix}\text{Size of}\\ \text{Sample}\\ \text{Space}\end{Bmatrix}$

k = number of steps in experiment

n_1 = number of possible outcomes in step 1

n_2 = number of possible outcomes in step 2

n_k = number of possible outcomes in Last Step

(b) If number of possible outcomes is the same each time:

$\begin{Bmatrix}\text{Total number of}\\ \text{Experimental outcomes}\\ \text{(sample points}\end{Bmatrix} = n^k$

(c) Examples:

① Experiment: Flip coin 3 times

$n_1 = 2$ (Heads or Tails)

$n_2 = 2$ (Heads or Tails

$n_3 = 2$ (Heads or Tails)

$k = 3$

$\begin{Bmatrix}\text{total \# of}\\ \text{experimental}\\ \text{outcomes}\end{Bmatrix} = \begin{Bmatrix}\text{Total \# of}\\ \text{sample}\\ \text{points}\end{Bmatrix} = 2*2*2 = 8$ or $2^3 = 8$

② Experiment: Roll die 2 times

$n_1 = 6$ (1, 2, 3, 4, 5, 6)

$n_2 = 6$ (1, 2, 3, 4, 5, 6)

$k = 2$

$\begin{Bmatrix}\text{Total \# of sample points}\\ \text{total \# of experimental outcomes}\end{Bmatrix} = 6*6 = 36$ or $6^2 = 36$

Remember from math:

$n! = n$ factorial

Example:

$4! = 4*3*2*1 = 24$

$$\frac{10!}{7!} = \frac{10*9*8*\cancel{7}*\cancel{6}*\cancel{5}*\cancel{4}*\cancel{3}*\cancel{2}*\cancel{1}}{\cancel{7}*\cancel{6}*\cancel{5}*\cancel{4}*\cancel{3}*\cancel{2}*\cancel{1}}$$

$$= 10*9*8 = 720$$

$$\frac{4!}{(4-4)!} = \frac{4!}{0!} = \frac{4*3*2*1}{1} = 24$$

$\boxed{0! = 1}$

In Excel use: FACT Function

7) count total number of experimental outcomes (sample points) when selecting n objects from a set of N objects (order does not matter, $(1,2) = (2,1)$) P6

---

a) $${}_N C_n = C^N_n = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \left\{ \begin{array}{l} \text{Counting Rule} \\ \text{for combinations} \end{array} \right\}$$

N = count of all objects (population size)

n = objects taken (items selected) (sample size)

b) <u>Example</u>:

Find all possible combinations of sample size n, from population with size N (used later with "Central Limit Theorem" ch. 7)

Boomerang Quad (Q)	population = $N = 4$
Boomerang Tri-Fly (T)	sample size = $n = 2$
Boomerang Sunshine (S)	
Boomerang Bellen (B)	

Randomly select 2 of 4 to test (order does not matter)

$$C^4_2 = \frac{4!}{2!(4-2)!} = \frac{4*3*\cancel{2}*\cancel{1}}{2*1 * \cancel{2}*\cancel{1}} = \frac{12}{2} = 6$$

Excel = COMBIN (number, number_chosen)
= COMBIN (N, n)
= COMBIN (4,2) = 6

$$\left\{ \begin{array}{c} \text{Sample} \\ \text{space} \end{array} \right\} = S = \{(Q,T), (Q,S), (Q,B), (T,S), (T,B), (S,B)\}$$

sample points Experimental outcomes

8 Counting Rule for Permutations $\left(\begin{array}{c}\text{order matters}\\ (2,1) \neq (1,2)\end{array}\right)$

a) $${}_N P_n = P^N_n = \binom{N}{n} = \frac{N!}{(N-n)!}$$

$N =$ count of all objects
$n =$ objects taken

b) Example:

Randomly select 2 of 4 boomerangs to test (order matters).

Boom Q
Boom T
Boom S
Boom B

$N = 4$
$n = 2$

$$P^4_2 = \frac{4!}{(4-2)!} = \frac{4!}{2!} = \frac{4*3*\cancel{2}*\cancel{1}}{\cancel{2}*\cancel{1}} = 12$$

Excel = PERMUT(number, number_chosen)
= PERMUT(N, n)
= PERMUT(4, 2) = 12

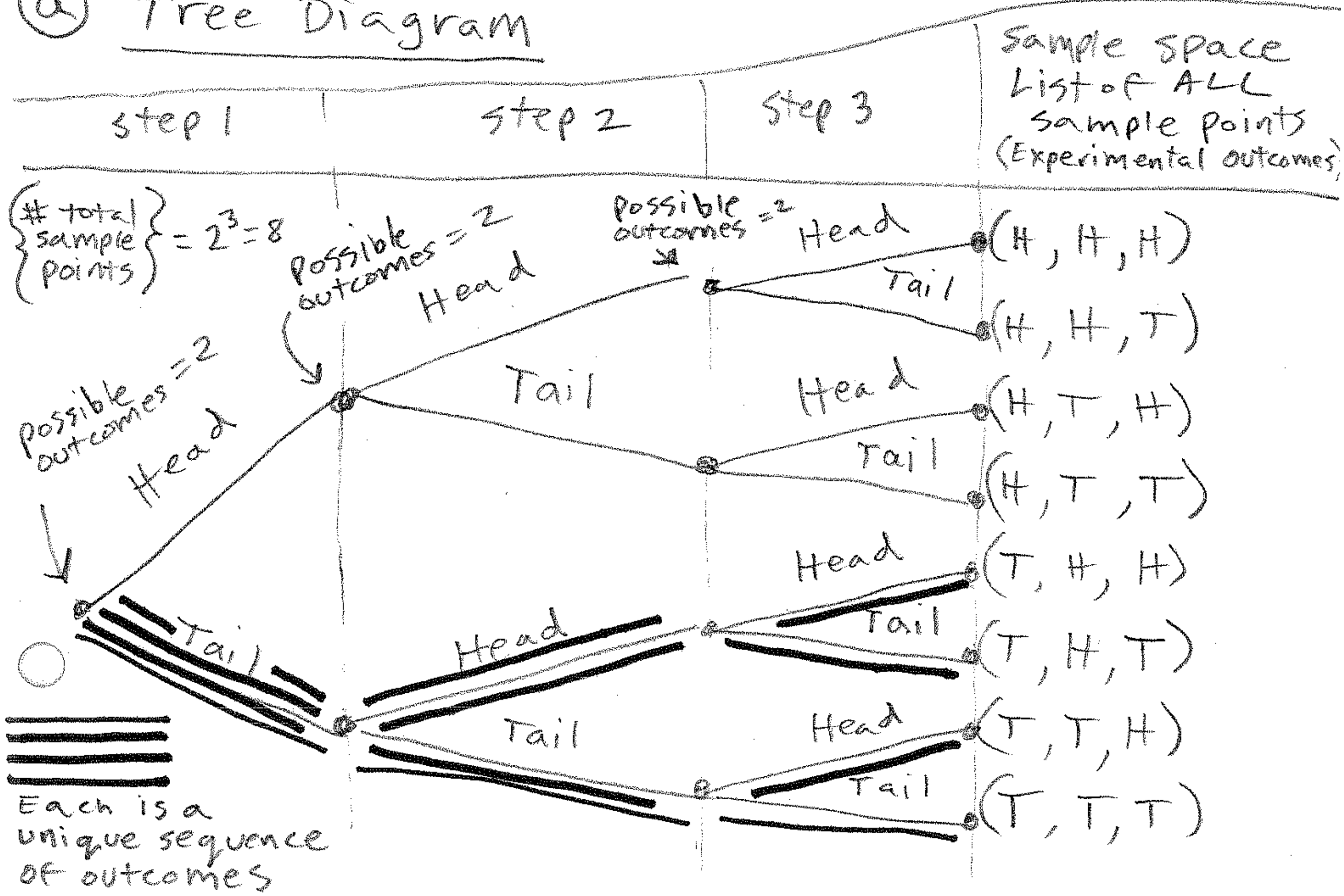
$\left\{\begin{array}{c}\text{Sample}\\ \text{Space}\end{array}\right\} = S = \left\{\begin{array}{l}(Q,T), (Q,S), (Q,B), (T,Q), (T,S), (T,B),\\ (S,Q), (S,T), (S,B), (B,Q), (B,T), (B,S)\end{array}\right\}$

Each is a Sample Point / Experimental outcome

(9) List ALL sample points in Sample Space (P8)

Multi-step Random Experiment: Toss Coin 3 Times

(a) Tree Diagram



(b) Table Format

	1st Toss	2nd Toss	3rd Toss	Sample Space ALL sample points
1	H	H	H	(H, H, H)
2	H	H	T	(H, H, T)
3	H	T	H	(H, T, H)
4	T	H	H	(T, H, H)
5	T	T	H	(T, T, H)
6	T	H	T	(T, H, T)
7	H	T	T	(H, T, T)
8	T	T	T	(T, T, T)

9) c Example:

Experiment: Roll 2 dice

$n_1 = 6$

$n_2 = 6$

$K = 2$

$\left\{ \begin{array}{l} \text{Total \# of} \\ \text{sample points} \end{array} \right\}$ $6 * 6 = 6^2 = 36$

sample space (List of sample points)

	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6

Methods of Assigning Probability

(a) Classical Method

All experimental outcomes are equally likely.

$$P(E_i) = \frac{\text{\# of favorable sample points}}{\text{total \# of possible sample points}}$$

Example: [1] Roll die, 6 equally likely outomes

$P(\text{Roll a } 4) = \frac{1}{6}$

"4" / 6 sides

[2] Get Audited by district office, 3000 out of 3,000,000 Tax Returns

$P(\text{Audit}) = \frac{3000}{3000000} = \frac{1}{1000}$

(b) Relative Frequency Method

we already saw this Method in Ch. 2

Use past Data to Predict Future

or

Use past Data Available to estimate the proportion of the time the experimental outcome will occur if the Experiment is repeated a large number of times.

Example: instructor gave 10 A's out of 100 in past class. $P(A) = \frac{10}{100} = \frac{1}{10}$

...arge numbers:
over a large number of Trials, Relative Frequency will approach True probability.

(c) Subjective Method

[1] Can't realistically assume outcomes equally likely

[2] Little Past Data exists

[3] You will use the best information you have, but it will be your personal belief.

Examples: who will win Super Bowl, merger between Google & Microsoft

Requirements for Assigning Probabilities

Prob. between 0 & 1

$\boxed{1}$ $$0 \leq P(E_i) \leq 1 \quad \text{for all } i$$

E_i = i^{th} Experimental outcome (sample Point)

$P(E_i)$ = Probability

sum of ALL = 1

$\boxed{2}$ $$\sum P(E_i) = 1$$

or

$$P(E_1) + P(E_2) + \dots + P(E_N) = 1$$

sum of all pobabilities for <u>ALL</u> experimental outcomes (sample Points) must be one!!

- ☆ We already saw these 2 requirements in action when we made <u>Relative Frequency</u> Tables!
- ☆ when using classical method, the 2 requirements are automatically satisfied. Dice example: $\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=1$
- ☆ when using Relative Frequency Method, the 2 Req. are automatically satisfied if categories are mutually exclusive & collectively exaustive.
- ☆ Because S, sample space, is an event, and it contains all sample points, $P(S) = \underline{1}$

Name	Hair Color	Eye Color
Sioux	Brown	Hazel
Chin	Black	Blue
Shelia	Black	Brown
Gigi	Brown	Hazel
Tyrone	Black	Brown
Lin	Blond	Brown
Ducky	Brown	Brown
Kip	Brown	Hazel
Linda	Blond	Blue
Jo	Blond	Green

Chapter 2:

Relative Frequency:

Raw Data

Relative Frequency

Hair color	Frequency	Relative Frequency
Brown	4	$4/10 = 0.40$
Black	3	$3/10 = 0.30$
Blond	3	$3/10 = 0.30$
total	10	$\sum = 1$

Rule 1 $\sum = 1$ ✓

Rule 2 each $P(E_i)$ is:

$$0 \le P(E_i) \le 1$$ ✓

12 Event

A collection of 1 or more sample points (Experimental outcomes)

13 Probability of an Event P(E)

The probability of any event is equal to the sum of the probabilities of the sample points in the event.

ⓐ **Example:**

K=3 Experiment: Toss coin 3 times

$n = 2$

$2^3 = 8 =$ total sample points

	Toss 1	Toss 2	Toss 3	List of Sample Points	Probability of each	Define success as Tail. Count Tails
1	H	H	H	H, H, H	1/8 = 0.125	0
2	H	H	T	H, H, T	1/8 = 0.125	1
3	H	T	H	H, T, H	1/8 = 0.125	1
4	T	H	H	T, H, H	1/8 = 0.125	1
5	T	T	H	T, T, H	1/8 = 0.125	2
6	T	H	T	T, H, T	1/8 = 0.125	2
7	H	T	T	H, T, T	1/8 = 0.125	2
8	T	T	T	T, T, T	1/8 = 0.125	3

A) Event = Get 2 Tails in 3 Flips = {3 sample points match this requirement}

Sample Points or Events = {(T, T, H), (T, H, T), (H, T, T)}

P(Get 2 "T" in 3 Flips) = P(T, T, H) 0.125 + P(T, H, T) 0.125 + P(H, T, T) 0.125

= 0.125 + 0.125 + 0.125 = 0.375

(13) (a)

(B) Event = Get 2 or more Tails in 3 Flips

Sample Points or Event = $\{(T,T,H),(T,H,T),(H,T,T),(T,T,T)\}$

$P(\text{Get 2 or More "T" in 3 Flips}) = \frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8} = \frac{4}{8} = \frac{1}{2}$
$= 0.5$

(c) Event = Get 3 straight Tails

sample points for event = $\{(T,T,T)\}$

$P(\text{Get 3 straight Tails}) = \frac{1}{8} = 0.125$

---

(13)

(b) Example:

Experiment: Toss 2 die

steps = K = 2
outcomes = n = 6
$6^2 = 36 =$ total sample points

	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6

$P(\text{Roll a } 7) = ?$

sample points or Event $= \{(6,1),(5,2),(4,3),(3,4),(2,5),(1,6)\}$

$P(\text{Roll 7 w/ 2 die}) = \frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36} = \frac{6}{36} = \frac{1}{6}$

(14) Venn Diagram

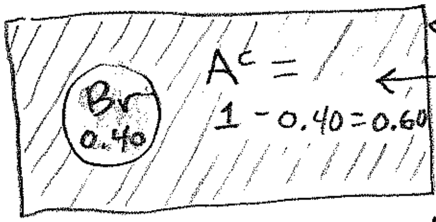
Diagrams to show relationships between different events and the sample space

Name	Hair Color	Eye Color
Sioux	Brown	Hazel
Chin	Black	Blue
Shelia	Black	Brown
Gigi	Brown	Hazel
Tyrone	Black	Brown
Lin	Blond	Brown
Ducky	Brown	Brown
Kip	Brown	Hazel
Linda	Blond	Blue
Jo	Blond	Green

(a) Complement

Event Br = Brown Hair

$P(Br) = \frac{4}{10} = 0.40$



← sample space = 1

← complement of Event A = A^c

$P(A) = 1 - P(A^c)$

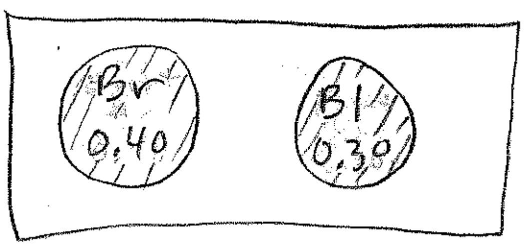
$P(A^c) = 1 - P(A)$

A^c = All sample points NOT A.

"Dating only one person"

(b) Mutually Exclusive

= 2 Events are Mutually Exclusive if they have NO sample points in common.



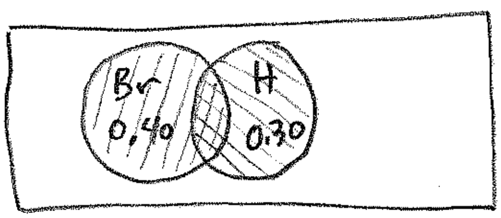
Event Br = Brown Hair, $P(Br) = 0.40$

Event Bl = Black Hair, $P(Bl) = \frac{3}{10} = 0.30$

(c) Union

= The union of Br & H is the Event containing all sample points for Br or H or Both.

OR



Event Br = Brown Hair, $P(Br) = 0.40$

Event H = Hazel Eyes, $P(H) = 0.30$

union = Br OR H OR Both = OR criteria

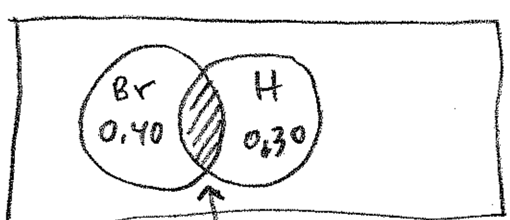
= $\cup$ = "At least 1" = "1 or more"

symbol ↑

(d) Intersection

= The intersection of Br & H is the Event containing the sample points for both Br AND H (2 TRUEs)

AND



Intersection = Br AND H = AND Criteria

= Both = Joint = Concurrent = 2 TRUEs

= $\cap$ ← symbol

Br AND H (2 TRUEs) = $\frac{3}{10}$ = 0.30

15 Addition Law for Probability a & b

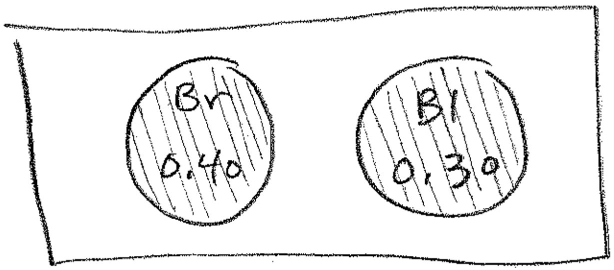
Name	Hair Color	Eye Color
Sioux	Brown	Hazel
Chin	Black	Blue
Shelia	Black	Brown
Gigi	Brown	Hazel
Tyrone	Black	Brown
Lin	Blond	Brown
Ducky	Brown	Brown
Kip	Brown	Hazel
Linda	Blond	Blue
Jo	Blond	Green

a Mutually Exclusive Events

$$P(Br) = \frac{4}{10} = 0.40$$

$$P(Bl) = \frac{3}{10} = 0.30$$

Add Rule



$$P(Br \text{ OR } Bl) = P(Br \cup Bl) = P(Br) + P(Bl)$$

$$= 0.40 + 0.40$$

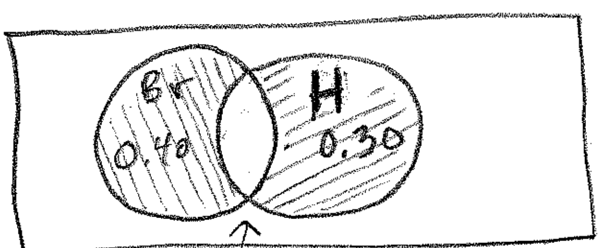
$$= 0.70$$

b Not Mutually Exclusive Events

$$P(Br) = \frac{4}{10} = 0.40$$

$$P(H) = \frac{3}{10} = 0.30$$

$$P(Br \text{ AND } H) = P(Br \cap H) = \frac{3}{10} = 0.30$$

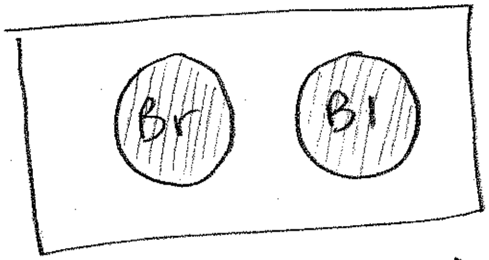


Must subtract so we don't Double count!!

$$P(Br \text{ OR } H) = P(Br \cup H) = P(Br) + P(H) - P(Br \text{ AND } H)$$

$$= 0.40 + 0.30 - 0.30 = 0.40$$

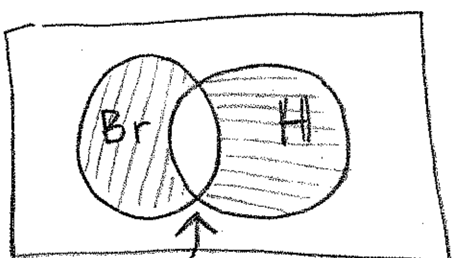
(c) Summary of Addition Law for Probability



{Mutually Exclusive Events}

$$P(Br \text{ OR } Bl) = P(Br) + P(Bl)$$

$$P(Br \cup Bl) = P(Br + P(Bl)$$



{NOT Mutually Exclusive Events}

$$P(Br \text{ OR } H) = P(Br) + P(H) - P(Br \text{ AND } H)$$

$$P(Br \cup H) = P(Br) + P(H) - P(Br \cap H)$$

{Must subtract so you DO NOT DOUBLE COUNT}

Example 1:

Flight Arrival		Frequency
Early	(E)	100
on time	(OT)	800
Late	(L)	75
canceled	(C)	25
		1000

Mutually Exclusive Events

$$P(E) = \frac{100}{1000} = .1$$

$$P(E \text{ or } OT) = P(E \cup OT) = \frac{(800+100)}{1000} = .9$$

Example 2:

Travel survey of People who visit Seattle

Visit Space Needle = SN

Visit Pike's Place Market = PP

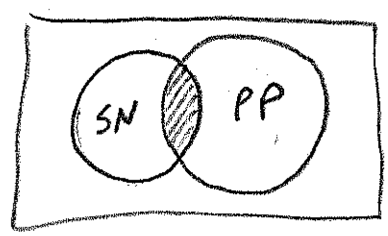
Not M.E. events

$P(SN) = .45$

$P(PP) = .65$

$P(SN \text{ and } PP) = .25 \quad = P(SN \cap PP)$

$P(SN \text{ or } PP) = .45 + .65 - .25$



MUST subtract so we don't count twice!!

* Note problem may say: P(at least one site)
P(Either site)

Example 3

Experiment asking 2 questions about Heart Attack & smoking of people over 50 (Random sample):

	Heart Attack?		
Smoke?	Yes	No	Total
Do not smoke	30	220	250
Moderate	60	65	125
Heavy	90	35	125
Total	180	320	500

Event = Yes had Heart Attack OR Heavy smoker

$$P(\text{Yes OR Heavy}) = P(\text{Yes}) + P(\text{Heavy}) - P(\text{Yes AND Heavy})$$

$$P(\text{Yes OR Heavy}) = \frac{180}{500} + \frac{125}{500} - \frac{90}{500}$$

$$= \frac{180 + 125 - 90}{500} = \frac{215}{500}$$

$$= \frac{43}{100} = 0.43$$

Creating AN nd OR Criteria Counting Formulas in Excel

	A	B	C	D	E	F	G	H	I	J	K
1	**Name**	**Hair Color**	**Eye Color**								
2	Sioux	Brown	Hazel								
3	Chin	Black	Blue								
4	Shelia	Black	Brown								
5	Gigi	Brown	Hazel								
6	Tyrone	Black	Brown								
7	Lin	Blond	Brown								
8	Ducky	Brown	Brown								
9	Kip	Brown	Hazel								
10	Linda	Blond	Blue								
11	Jo	Blond	Green								
12											
13	**AND Criteria**										
14											
15	**Hair Color**	**Eye Color**	**Count**								
16	Brown	Hazel	3	=COUNTIFS(B2:B11,A16,C2:C11,B16)							
17											
18	**OR Criteria**										
19											
20	**Hair Color**	**Eye Color**	**Count**								
21	Brown	Hazel	4	=COUNTIFS(B2:B11,A21)+COUNTIFS(C2:C11,B21)-COUNTIFS(B2:B11,A21,C2:C11,B21)							
22			(Count of Brown Hair Color) + (Count of Hazel Eye Color) - (Count of Brown Hair Color AND Brown Hair Color)								

subtract so you don't Double count!!

P 20

Conditional Probability

(a) The probability of an event given that another event already occurred.

(b) After the first event has occurred the "sample space has changed".

(c) Notation:

$P(A \mid B)$ "Probabity of A given that B has already occurred"

"Given That"

(d) Examples:

(1) probability of pulling a Queen from a deck of cards (without replacement) affects the probability of pulling the next card.

Event Q_1 = Pull Queen from 52 card deck
Event Q_2 = Pull 2nd Queen AFTER 1 Queen Pulled

$P(Q_1) = \frac{4}{52}$

$P(Q_2 \mid Q_1) = \frac{3}{51}$ ← {51 cards left. Sample Space has changed}

(2) Pulling female name from hat filled with 18 female names & 13 male names affects next pull. $P(F_1) = \frac{18}{31}$ $P(F_2 \mid F_1) = \frac{17}{30}$

(3) $P\left(\text{Heavy Smoker} \mid \text{over 50 and have not had a Heart Attack}\right) = \frac{35}{320}$

4) Cross Tabulated Tables

Each intersecting cell is counting with 2 criteria or 2 Events

Random Sample of 200 Purchases at Convenience Store

		Amount				
		$0 up to $25	$25 up to $50	$50 up to $75	>= &75	Total
Payment Method	Cash	21	11	9	3	**44**
	Debit	20	15	14	8	**57**
	Credit	16	28	27	15	**86**
	Check	6	3	2	2	**13**
	Total	**63**	**57**	**52**	**28**	**200**

$$P(\text{cash} \mid 0\text{-}25) = \frac{21}{63} = \frac{1}{3} = 0.\overline{3}$$

5) Cross Tabulated Table with % of Grand Total (Joint Probability Table)

"Joint" "Intersecting" "AND" "Concurrent" Probabilities

		Amount				
		$0 up to $25	$25 up to $50	$50 up to $75	>= &75	Total
Payment Method	Cash	10.5%	5.5%	4.5%	1.5%	**22.0%**
	Debit	10.0%	7.5%	7.0%	4.0%	**28.5%**
	Credit	8.0%	14.0%	13.5%	7.5%	**43.0%**
	Check	3.0%	1.5%	1.0%	1.0%	**6.5%**
	Total	**31.5%**	**28.5%**	**26.0%**	**14.0%**	**100.0%**

Marginal Probabilities

$$P(\text{cash} \mid 0\text{-}25) = \frac{10.5\%}{31.5\%} = \frac{0.105}{0.315} = \frac{1}{3} = 0.\overline{3}$$

Notice → $\dfrac{P(0\text{-}25 \text{ AND cash})}{P(0\text{-}25)} = P(\text{cash} \mid 0\text{-}25)$

		Amount				
		$0 up to $25	$25 up to $50	$50 up to $75	>= &75	Total
Payment Method	Cash	33.3%	19.3%	17.3%	10.7%	**22.0%**
	Debit	31.7%	26.3%	26.9%	28.6%	**28.5%**
	Credit	25.4%	49.1%	51.9%	53.6%	**43.0%**
	Check	9.5%	5.3%	3.8%	7.1%	**6.5%**
	Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**

6) Cross Tabulated Table with % of Column Totals

$$P(\text{cash} \mid 0\text{-}25) = 33.3\% = 0.\overline{3}$$

(a) Each one of the columns (Not Total) are conditional probabilities

(7) Look at Joint Probability Table (P 23)

{Joint Probability Table} = {Cross Tabulated with % of Grand Totals}

		Amount				
		$0 up to $25	$25 up to $50	$50 up to $75	>= &75	Total
Payment Method	Cash	10.5%	5.5%	4.5%	1.5%	**22.0%**
	Debit	10.0%	7.5%	7.0%	4.0%	**28.5%**
	Credit	8.0%	14.0%	13.5%	7.5%	**43.0%**
	Check	3.0%	1.5%	1.0%	1.0%	**6.5%**
	Total	**31.5%**	**28.5%**	**26.0%**	**14.0%**	**100.0%**

"Joint"
"AND"
"Intersecting"
"concurrent"
Probabilities

Conditional probability

P(cash | 0-25) =

$\frac{10.5\%}{31.5\%} = \frac{\text{Joint}}{\text{Marginal}}$

Marginal Probabilities

Values in "Margins" that show Probabilities of each event separately

Conditional Probability Rule

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

or

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

conditional "Joint" "AND" Marginal

(8) Tree Diagrams (Probability Trees)

"Joint" "Intersection" "AND" Probabilities

Conditional Probabilities

Marginal Probabilities

Heart Attack → Smoke?

- No $\frac{320}{500}$
  - $\frac{220}{320}$ No → $\frac{320}{500} \cdot \frac{220}{320} = \frac{220}{500} = 0.44$
  - $\frac{65}{320}$ moderate → $\frac{320}{500} \cdot \frac{65}{320} = \frac{65}{500} = 0.13$
  - $\frac{35}{320}$ Heavy → $\frac{320}{500} \cdot \frac{35}{320} = \frac{35}{500} = 0.0$
- Yes $\frac{180}{500}$
  - $\frac{30}{180}$ No → $\frac{180}{500} \cdot \frac{30}{180} = \frac{30}{500} = 0.06$
  - $\frac{60}{180}$ Moderate → $\frac{180}{500} \cdot \frac{60}{180} = 0.12$
  - $\frac{90}{180}$ Heavy → $\frac{180}{500} \cdot \frac{90}{180} = 0.18$

$$P\left(\text{Heavy Smoker} \mid \text{Did Not have a Heart Attack}\right) = \frac{35}{320} = 0.109375$$

(9) In Excel, use HYPGEOMDIST Function for Conditional Probabilities → More later

Independent Events

(a) 2 events are independent if the probability of one event is not affective by the occurrence of the other event.

(b) Rule of Independence:

$P(A|B) = P(A)$ B has No affect on A

or

$P(B|A) = P(B)$ A has No affect on B

otherwise the events are dependent such as: $(P(Q_1) = \frac{4}{52}, P(Q_2|Q_1) = \frac{3}{51})$

(c) Examples of Independent Events:

1. Rolling 6 with die, does not affect Next roll.
2. A Traffic jam today, generally does not affect whether you are in traffic jam tomorrow.
3. whether or not Google stock goes up does not affect whether or not whole Food's stock goes up.

(18) Multiplication Law for Probability

(a) $P(A \text{ AND } B) = P(A \cap B) = P(A) * P(B \mid A)$

OR

$P(A \text{ AND } B) = P(A \cap B) = P(B) * P(A \mid B)$

(b) Multiplication Law for Independent Events:

$P(A \text{ AND } B) = P(A \cap B) = P(A) * P(B)$

since: $P(B) = P(B \mid A)$

$P(A) = P(A \mid B)$

Examples $\longrightarrow$

Notes: (1) AND $= \cap = * =$ Multiply

OR $= \cup = + =$ Add

(2) when checking for Independence you can use either: $P(A) = P(A \mid B)$

or

$P(B) = P(B \mid A)$

or

$P(A \text{ AND } B) = P(A \cap B) = P(A) * P(B)$

(3) Don't confuse "Mutually Exclusive" (No sample points in common; if one event occurs, the other did not) and "Independence" (The 2 events exist, but are not related).

Multiplication Law Examples

(1) The probability that Google stock will go up next year is estimated to be 0.35 & the probability that Whole Foods stock will go up is estimated to be 0.20. What is Probability that both will go up?

Assume events are independent

$P(G) = 0.35$

$P(WF) = 0.20$

$P(G \text{ AND } WF) = 0.35 * 0.20$
$= 0.07$

(2) Probability that product coming off assembly line is defected is estimated to be 0.02. What is probability that 3 straight would be defective?

$P\left(\begin{array}{l}\text{Randomly selecting 3} \\ \text{Products \& all 3 defective}\end{array}\right) = 0.02 * 0.02 * 0.02$
$= 0.000008$

very small, so if three are defective, there is a problem.

(3) What is Probability that you pull 2 straight Queens from a deck of cards?

$P(Q_1 \text{ AND } Q_2) = P(Q_1) * P(Q_2 \mid Q_1) = \frac{4}{52} * \frac{3}{51} = 0.0045$

(4) Stats class:

18 women
13 men

31 Total

Conditional

$$P\left(\begin{array}{c}\text{select 5 straight women}\\ \text{names from Hat}\end{array}\right) = \frac{18}{31} * \frac{17}{30} * \frac{16}{29} * \frac{15}{28} * \frac{14}{27}$$

$$= 0.0504264$$

"Joint" "AND"

In Excel:

success = pull woman name

probability mass function

$$= \text{HYPGEOM.DIST}(5, 5, 18, 31, 0)$$

# of successes

sample size

successes in population

population size

(5) Visit Seattle

P(visit SN) = 0.45 = Space Needle
P(visit PP) = 0.65 = Pike's Place Market
P(SN & PP) = 0.25 = Both

Find $P(SN \mid PP) = \dfrac{P(SN \text{ AND } PP)}{P(PP)} = \dfrac{0.25}{0.65} = 0.3846$

"Given that"

Q 29

Tree Diagram

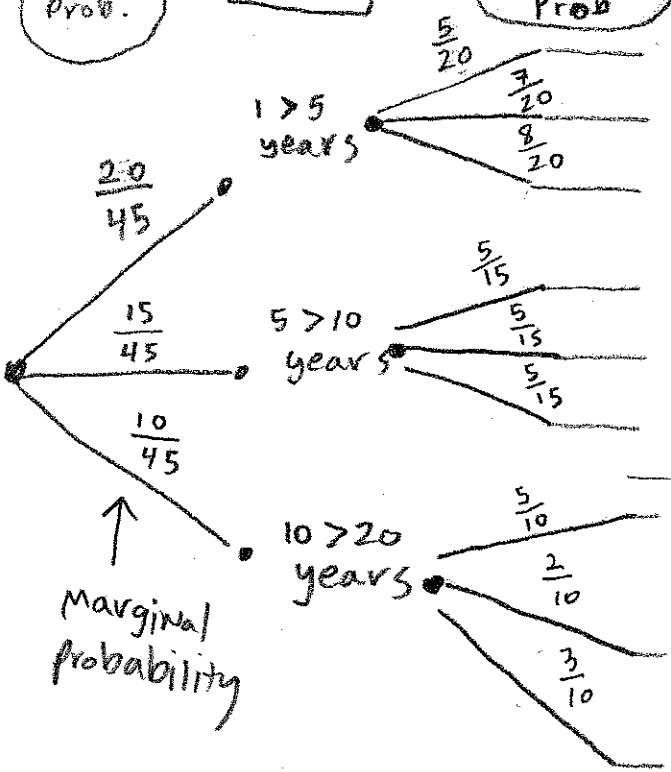
shows all conditional and joint probabilities for a Cross Tab table

1. Heavy dot represents root
2. Draw branches with Root Prob.
3. Draw branches with conditional Prob.
4. Write out Joint Probabilities

Tenure and Class ratings

Class rating	Had Tenure for # of years: 1 > 5 years	5 > 10 years	10 > 20 years	Total
OK	5	5	5	15
Good	7	5	2	14
Excellent	8	5	3	16
Total	20	15	10	45

Root Prob. | Tenure | Conditional Prob | Class Rating | Joint Probabilities



OK → $\frac{20}{45} * \frac{5}{20} = \frac{5}{45} = \frac{1}{9}$

Good → $\frac{20}{45} * \frac{7}{20} = \frac{7}{45} = \frac{7}{45}$

Excellent → $\frac{20}{45} * \frac{8}{20} = \frac{8}{45} = \frac{8}{45}$

OK → $\frac{15}{45} * \frac{5}{15} = \frac{5}{45} = \frac{1}{9}$

Good → $\frac{15}{45} * \frac{5}{15} = \frac{5}{45} = \frac{1}{9}$

Excellent → $\frac{15}{45} * \frac{5}{15} = \frac{5}{45} = \frac{1}{9}$

OK → $\frac{10}{45} * \frac{5}{10} = \frac{5}{45} = \frac{1}{9}$

Good → $\frac{10}{45} * \frac{2}{10} = \frac{2}{45} = \frac{2}{45}$

Excellent → $\frac{10}{45} * \frac{3}{10} = \frac{3}{45} = \frac{1}{15}$

Problem 1

A deli offers 5 cheeses, 4 meats and 3 rolls. How many different sandwich combinations are possible

n = # of possible outcomes

n_1 = # of cheeses = 5
n_2 = # of meats = 4
n_3 = # of rolls = 3

counting Rule:

$$\left\{\begin{matrix}\text{total number} \\ \text{of experimental} \\ \text{outcomes} \\ \text{(sample points)}\end{matrix}\right\} = n_1 * n_2 * \ldots n_k = \left\{\begin{matrix}\text{size of} \\ \text{sample} \\ \text{space}\end{matrix}\right\}$$

$$\left\{\begin{matrix}\text{\# different} \\ \text{sandwich} \\ \text{combos}\end{matrix}\right\} = 5 * 4 * 3 = 60$$

Problem 2

3 scholarships are available for needy students at \$5000, \$7000 and \$10,000. 12 students have applied & No student may receive more than one scholarship. If all students are in need of funds, how many different ways could the scholarship be awarded? Order does not matter

N = count of all items (population size)
n = items taken (sample size)

$$_NP_n = P^N_n \binom{N}{n} = \frac{N!}{(N-n)!}$$

N = All students = 12
n = students awarded = 3

$$\left\{\begin{matrix}\text{\# ways awards} \\ \text{could be granted}\end{matrix}\right\} = \frac{12!}{(12-3)!} = \frac{12!}{9!} = 12 * 10 * 11 = 1320$$

Problem 3

Basketball team A lost 40 consecutive games, coach wanted to pick a team randomly (to shake things up) by selecting 5 names from 12 in a hat. Assuming a player can play any position, how many combos are possible?

$$_N C_n = C^N_c = \binom{N}{n} = \frac{N!}{n!\,(N-n)!}$$

N = 12 Team members

n = 5 on team

$$\left\{\begin{matrix}\text{\# of combos} \\ \text{for team}\end{matrix}\right\} = \frac{12!}{5!\,(12-5)!} = \frac{12!}{5!\,7!} = \frac{12*11*\cancel{10}^{2}*\cancel{9}^{3}*\cancel{8}^{4}}{\cancel{5}*\cancel{4}*\cancel{3}*\cancel{2}*1}$$

$$= 12*11*2*3*4$$

$$= 792$$

P 4

	Heart Attack Yes	No	
Dont Smoke	30	220	250
moderate smoker	60	65	125
Heavy smoker	90	35	125
	180	320	500

$$P\left(\text{Yes H.A. OR Heavy Smoke}\right) = P(YHA) + P(HS) - P(YHA \text{ AND } HS) =$$

$$= \frac{180+125-90}{500} = \frac{215}{500} = \frac{43}{100} = 0.43$$

$$P(HS \text{ AND No H.A}) = \frac{35}{500} = \frac{7}{100} = 0.07$$

P5

Amount of purchase

	<20	20=<P<50	>=50	
cash	15	10	5	30
check	10	30	20	60
charge	10	20	20	50
	35	60	45	140

$P(\text{cash or} <20) = P(\text{cash}) + P(<20) - P(\text{cash AND} <20)$

$= \frac{30}{140} + \frac{35}{140} - \frac{15}{140} = \frac{50}{140} = \frac{5}{14}$

$P(\text{check AND} >50) = P(\text{check}) * P(>50 \mid \text{check})$

$= \frac{60}{140} * \frac{20}{60} = \frac{20}{140} \; \frac{1}{7}$

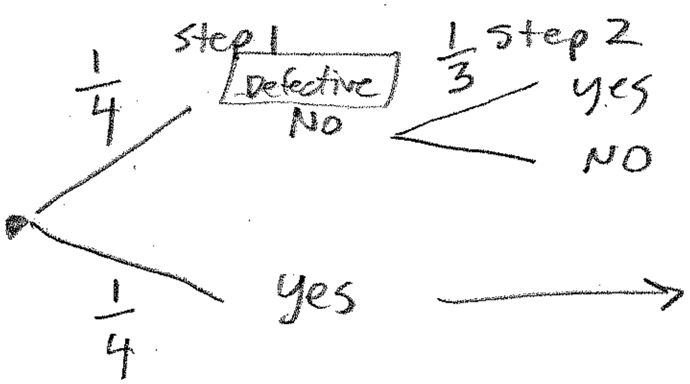
P6

shipment 4 carburetors
1 known to be defective
If 2 are selected at random & tested

① $P(\text{0 are defective}) = \frac{1}{4} * \frac{1}{3} = \frac{1}{12}$

② $P(\text{defect is found in 2 Tries}) = \frac{1}{4} * \frac{1}{3} + \frac{1}{4} = \frac{1}{12} + \frac{1}{4} = \frac{1}{12} + \frac{3}{12} = \frac{4}{12}$

$= \frac{1}{3}$



$\rightarrow \frac{1}{4} * \frac{1}{3} = \frac{1}{12}$

$\frac{1}{4} = \frac{1}{4} = \frac{3}{12}$

$+ \quad \frac{4}{12} = \frac{1}{3}$

P7

P.33

# of Doctor visits per Year	Frequency (men who go to Doctors)	Relative Frequency
0	30	$\frac{30}{300} = \frac{1}{10} = 0.10$
1	60	$\frac{60}{300} = \frac{6}{30} = \frac{1}{5}$ 0.20
2	90	$90/300$ $\frac{9}{30} = \frac{3}{10} = .30$
3 or more	120	$120/300 = \frac{12}{30} = \frac{4}{10} = .40$
total	300	$300/300 = 1$

each is ✓
$0 \le P \le 1$

$\Sigma = 1$ ✓

$P(2) = 0.30$

$P(\text{visit Doctor}) = 1 - 0.10 = 0.90$

$P(\text{2 or more}) = .3 + .4 = .70$

P8

$P(\text{Read MT}) = 0.45$
$P(\text{Read UT}) = 0.60$
$P(\text{Both}) = 0.20$

$P(\text{MT or UT}) = P(\text{Read at least } 1) = P(\text{MT, UT, or both}) = .45 + .6 - .2 = 0.85$

P.9

# cans sold	# of Days	Relative Frequency
0	5	$\frac{5}{60} = \frac{1}{12} = 0.8\overline{3}$
1	5	$\frac{5}{60} = \frac{1}{12} = 0.8\overline{3}$
2	10	$\frac{10}{60} = \frac{1}{6} = 0.1\overline{6}$
3	20	$\frac{20}{60} = \frac{1}{3} = 0.\overline{3}$
4	15	$\frac{15}{60} = \frac{1}{4} = 0.25$
>=5	5	$\frac{5}{60} = 0.8\overline{3}$
total	60	$\Sigma = 1 = \frac{12}{12}$ ✓

$0 \le P \le 1$ ✓

$$\begin{array}{r} 0.0833 \\ 12\overline{)1.000} \\ 960 \\ 40 \\ 36 \\ 40 \end{array}$$

P.34

$$P(2) = \frac{1}{6}$$

$$P(3 \text{ or more}) = \frac{20+15+5}{60} \quad \frac{40}{60} = \frac{4}{6} = \frac{2}{3}$$

$$P(\text{at least } 1) = \frac{12}{12} - \frac{1}{12} = \frac{11}{12}$$

20

Bayes' Theorem

Used when we have initial probabilities, called Prior Probabilities, we are given or get new information, and we want to revise or update the prior probabilities by calculating Posterior Probabilities.

Prior probabilities (subjective) → New information (Relative Frequency) → Apply Bayes' Theorem → Posterior (Revised) probabilities

You estimate the probability that you can get a passing CPA score of 750 or more points as

$P(\geq 750) = 0.25$

New info = survey Report that asked people who took CPA:

1. what was their score
2. Did they take preparation course

Report said:

$P(PC \mid \geq 750) = 0.60$

$P(PC \mid < 750) = 0.31$

$$\frac{P(\geq 750 \text{ AND } PC)}{\left(P(\geq 750 \text{ AND } PC) + P(< 750 \text{ AND } PC) \right)}$$

$$\frac{0.15}{0.15 + 0.2325}$$

$P(\geq 750 \mid PC) = 0.39$

Revised

Revised

Posterior Probabilities

Prior Probabilities

Event = CPA score ≥ 750 $\Rightarrow P(\geq 750) = 0.25$

Event = CPA Score < 750 $\Rightarrow P(< 750) = 1 - 0.25 = 0.75$

Events are: Mutually Exclusive & collectively exhaustive

These are your estimates

↑ subjective

Event = took preparation course = $P(PC)$

Event = NOT take prep. course = $P(NPC)$

Relative Frequency

New information from survey

$P(PC \mid \geq 750) = 0.60$

$P(PC \mid < 750) = 0.31$

Not add to 1
Not list of all conditional prob.

→ see Probability Tree Next Page

see Probability Tree Next Page

Prior Probabilities	Conditional Probabilities	Joint/AND Probabilities	Posterior Probabilities
$P(\geq 750) = 0.25$	$P(PC \mid \geq 750) = 0.6$	$P(\geq 750 \text{ AND } PC) = 0.25 * 0.6 = 0.15$	$P(\geq 750 \mid PC) = \frac{0.15}{0.3825} = 0.39$
$P(< 750) = 0.75$	$P(PC \mid < 750) = 0.31$	$P(< 750 \text{ AND } PC) = 0.75 * 0.31 = 0.2325$	$P(< 750 \mid PC) = \frac{0.2325}{0.3825} = 0.61$

$$\Sigma = \begin{array}{r} 0.15 \\ +0.2325 \\ \hline 0.3825 \end{array}$$

↑ Whole = $P(PC)$

↑ compare parts to whole

* As long as Events "≥ 750" & "< 750" mutually exclusive & collectively Exhaustive

Conclusion: with new info we see that Probability that we pass CPA given that we take prep. course is 0.39. Better than our original of 0.25

Joint / AND Probabilities	Posterior Probabilities
$P(\geq 750 \text{ AND } Pc) = 0.15$	$P(\geq 750 \mid Pc) = \frac{0.15}{0.3825} = 0.39$
$P(< 750 \text{ AND } Pc) = 0.2325$	$P(< 750 \mid Pc) = \frac{0.2325}{0.3825} = 0.61$
$P(Pc) = 0.3825$	

Multiplication Law & Conditional Probability Formula:

$$\boxed{\text{conditional} = \frac{\text{AND / Joint}}{\text{Marginal}}}$$

$$P(\geq 750 \mid Pc) = \frac{P(\geq 750 \text{ AND } Pc)}{P(Pc)}$$

$$P(\geq 750 \mid Pc) = \frac{P(\geq 750 \text{ AND } Pc)}{P(\geq 750 \text{ AND } Pc) + P(< 750 \text{ AND } Pc)}$$

Bayes' Theorem Formula (2 Event case)

Prior Event 1 = A_1

Prior Event 2 = A_2

Event = B

$$P(A_1 \mid B) = \frac{P(A_1 \text{ AND } B)}{P(A_1 \text{ AND } B) + P(A_2 \text{ AND } B)}$$

$$P(A_2 \mid B) = \frac{P(A_2 \text{ AND } B)}{P(A_2 \text{ AND } B) + P(A_1 \text{ AND } B)}$$

or

$$P(A_1 \mid B) = \frac{P(A_1) * P(B \mid A_1)}{P(A_1) * P(B \mid A_1) + P(A_2) * P(B \mid A_2)}$$

$$P(A_2 \mid B) = \frac{P(A_2) * P(B \mid A_2)}{P(A_2) * P(B \mid A_2) + P(A_1) * P(B \mid A_1)}$$

Tree Diagram (Probability Tree)

Prior	Conditional	Joint AND
$P(\geq 750) = 0.25$	$P(PC \mid \geq 750) = 0.6$	$P(\geq 750 \text{ AND } PC) = .25 * .6 = 0.15$
	$P(NPC \mid \geq 750) = 0.4$	$P(\geq 750 \text{ AND } NPC) = .25 * .4 = 0.1$
$P(< 750) = 0.75$	$P(PC \mid < 750) = 0.31$	$P(< 750 \text{ AND } PC) = .75 * .31 = 0.2325$
	$P(NPC \mid < 750) = 0.69$	$P(< 750 \text{ AND } NPC) = .75 * .69 = .5175$

Add in Denominator

$$P(\geq 750 \mid PC) = \frac{P(\geq 750 \text{ AND } PC)}{P(\geq 750 \text{ AND } PC) + P(< 750 \text{ AND } PC)} = \frac{0.15}{0.15 + 0.2325} = 0.39$$

Bayes' Theorem Example 2

prior probability: Event = Leave Garage Door open $\Rightarrow P(GO) = 0.20$

conditional probabilities:

- Event: stuff stolen given Garage open $\Rightarrow P(SS|GO) = 0.05$
- Event: Stuff stolen given Garage NOT open $\Rightarrow P(SS|GNO) = 0.01$

We want to know what is probability that Garage was open given stuff stolen $P(GO|SS)$

Prior	Conditional	Joint AND	Posterior
$P(GO)$ 0.20	$P(SS\|GO)$ 0.05	$P(GO \text{ AND } SS)$ $0.2 * 0.05$ $= 0.01$	$P(GO\|SS) = \frac{0.01}{0.018}$ $= 0.556$
$P(GNO)$ $1 - 0.20$ $= 0.80$	$P(SS\|GNO)$ 0.01	$P(GNO \text{ AND } SS)$ $0.8 * 0.01$ $= 0.008$	$P(GNO\|SS) = \frac{0.008}{0.018}$ $= 0.444$
		$\Sigma = 0.01 + 0.008 = 0.018$	

whole → 0.018

parts compared to whole ↑

$P(SS) = 0.018$

therefor

$P(\text{Not } SS) = 1 - 0.018 = 0.982$