What is independence?

- Repeated independent trials
- The outcome of any trial of the experiment does not influence or affect the outcome of any other trial
- The trials are said to be physically independent
- Physical independence is a belief
- It cannot be proved that the trials are independent; we can only believe

Simple vs Compound Experiments

- Consider a simple experiment with sample space $\Omega = \{a_1, a_2, \ldots\}$
- The result of repeated independent trials of this experiment is a sequence or vector of outcomes, say, $(a_5, a_2, a_7, a_9, a_1, \ldots)$
- This vector is regarded as the outcome of a compound experiment with sample space $\Omega \times \Omega \times \Omega \times \ldots$
- Simple experiments are subexperiments

Events on Compound Experiments

- The outcome of a compound experiment is a sequence or vector of outcomes of the form

$$(a_5, a_2, a_7, a_9, a_1, \ldots)$$

- The simple event A occurred on i-th subexperiment if the i-th outcome in this sequence is a member of the event A Ω
- The compound event $(A, B, C, A^c, \ldots)$ occurred if a_5 A, a_2 B, a_7 C, a_9 A^c, …

Declaration of Independence

- The belief in independence is reflected in the assignment of probabilities to the events of the compound experiment
- If the trials are (believed to be) independent, then we set

  $P(A, B, C, A^c, \ldots) = P(A)P(B)P(C)P(A^c)\ldots$
- Both A and A^c cannot occur on the same trial of the simple experiment: here they are occurring on different subexperiments

What is the event A?

- We defined the event A on the sample space of the simple experiment
- The occurrence of A on the i-th trial can be viewed as an event A_i defined on the compound experiment
- Which outcomes of the compound experiment comprise A_i?
- All outcomes of the form

$$(✶, ✶, \ldots, x, ✶, ✶, \ldots)$$

  where x A and ✶ means "don't care"

Independence of the A_i

- Consider arbitrary events A_i and B_j defined on the compound experiment; $i \neq j$
- $B = A$ and even $B = A^c$ are acceptable choices as long as $i \neq j$
- Because of the physical independence of the subexperiments, we have that

$$P(A_i \cap B_j) = P(A_i)P(B_j)$$

- More generally,

  $P(A_1 \cap B_2 \cap C_3 \ldots) = P(A_1)P(B_2)P(C_3) \ldots$

More generally, …

- We apply this idea to general experiments
- Definition: Events A and B defined on an experiment are said to be (stochastically) mutually independent if

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

- Sometimes people say "A is independent of B" instead, but independence is mutual: A is independent of B if and only if B is independent of A

What's with the "stochastically"?

- Definition: Events A and B defined on an experiment are said to be (stochastically) mutually independent if

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

- If we believe that events A and B are physically independent, then we insist that this equality holds
- But, this equality can hold even for events that are provably physically dependent… the events are stochastically independent

Let's get physical… O. Newton-John

- Physical independence is, in essence, a property of the events themselves
- We believe that events A and B are physically independent and express this independence via P(AB) = P(A)P(B)
- Stochastic independence is a property of the probability measure
- Stochastic independence does not necessarily mean that the events are physically independent

I mean what I say …

- Physical independence (which is a belief, remember?) of A and B implies stochastic independence — we insist that we must have P(AB) = P(A)P(B)
- But, if we do not have any reason to believe that A and B are physically independent, and our calculations reveal that P(AB) = P(A)P(B), we should not automatically assume that A and B are also physically independent

… but I won't say what I mean!

- I hope that the difference between the notions of physical independence and stochastic independence will be retained in your mind
- In the future, we shall be using the word stochastically in conjunction with the word independent only on rare occasions
- Insist on P(AB) = P(A)P(B) where appropriate; but don't read too much into its serendipitious occurrence

Consequences of independence

- If A and B are mutually independent events, then P(AB) = P(A)P(B)
- This is equivalent to each of the following
  - $P(AB^c) = P(A)P(B^c)$
  - $P(A^cB) = P(A^c)P(B)$
  - $P(A^cB^c) = P(A^c)P(B^c)$
- In other words, A and B^c are mutually independent, as are A^c and B, and as for A^c and B^c, why, they are independent too!

A and B^c are mutually independent

- If A and B are mutually independent events, then P(AB) = P(A)P(B)
- But P(A) = P(AB) + P(AB^c) in general
- For independent events A and B, we get P(A) = P(A)P(B) + P(AB^c)
- Hence, P(AB^c) = P(A) – P(A)P(B)
  = P(A)(1 – P(B))
  = P(A)P(B^c)
- Similarly for the other two cases

A great comfort …

- Independence of A and B allows for easy computation of P(AB)
- It is a great temptation to apply it wherever possible
- Example: A and B are events with P(A) = 0.5, P(B) = 0.6 and P(A B) = 0.7. What is P(AB)?
- ~~Assuming A and B are independent, we get P(AB) = P(A)P(B) = 0.3~~

Actually, P(AB) = P(A)+P(B)–P(A B) = 0.4

Mutually is as mutually does…

- If A and B are mutually independent events, then P(AB) = P(A)P(B)
- If A and B are mutually exclusive events, then P(AB) = 0
- Do NOT confuse the two concepts
- Mutually exclusive events cannot be mutually independent (or vice versa) except in the trivial case when at least one of the two events A and B has zero probability

Conditional = unconditional

- If A and B are mutually independent events, then P(AB) = P(A)P(B)
- If P(A) > 0, we get that

  P(B|A) = P(AB)/P(A) = P(B)
- The conditional probability of B given A is the same as the unconditional probability!
- Knowing that A occurred does not cause any “updating” of the chances of B
- Similarly, P(A|B) = P(AB)/P(B) = P(A)

Saying it over and over…

- If A and B are mutually independent events, then P(AB) = P(A)P(B)
- If A and B are mutually exclusive events, then P(AB) = 0
- For mutually exclusive events, P(B|A) = 0
- Knowing that A occurred guarantees that B did not occur!
- Thus, A and B cannot be mutually independent as well

Why not the following definition?

- Many people (and textbook authors!) feel that

  P(B|A) = P(B)

  is a much more natural definition of the notion of independence
- “B is independent of A if P(B|A) = P(B)”
  - A and B seem to have different roles
  - Mutuality of independence is not obvious
  - Assumes that P(A) > 0

Exclusive-OR gates

- Example: Let A and B respectively denote the events that inputs #1 and #2 of an Exclusive-OR gate are logical 1

  Assume that A and B are physically independent (hence they are stochastically independent) events

  Assume that P(A) = P(B) = 0.5

  Let C denote the event that the output of the Exclusive-OR gate is logical 1
- $C = A \quad B = AB^c \quad A^cB$

Output depends on input?

- $P(A) = P(B) = 0.5;\ C = A \quad B = AB^c \quad A^cB$
- $P(C) = P(AB^c) + P(A^cB)$ Why?

  $= P(A)P(B^c) + P(A^c)P(B)$

  $= 0.5\times0.5 + 0.5\times0.5 = 0.5$
- Are A and C independent events?
- $P(AC) = P(A(AB^c \quad A^cB))$

  $= P(AB^c) = P(A)P(B^c)$

  $= 0.5\times0.5 = 0.25$

  $= P(A)P(C)$!!!!

XOR or not?

- Is the output of the XOR gate really independent of the input?
- The output is stochastically independent of the input
- The output is physically <u>dependent</u> on the input
- Physical independence (such as A and B being independent) is a belief
- Stochastic independence is an artifact of the probability measure

Repeat XOR gate example

- Example: Let A and B respectively denote the events that inputs #1 and #2 of an Exclusive-OR gate are logical 1

  Assume that A and B are physically independent (hence they are stochastically independent) events

  Assume that P(A) = P(B) = 0.500001

  Let C denote the event that the output of the Exclusive-OR gate is logical 1
- $C = A \quad B = AB^c \quad A^cB$

Output independent of input?

- $P(A) = P(B) = 0.500001$
- $C = A \quad B = AB^c \quad A^cB$
- $P(C) = 2\times0.500001\times0.499999$

  $= 0.49999999998$
- $P(AC) = P(AB^c) = P(A)P(B^c)$

  $= 0.500001\times0.499999$

  $= 0.2499999999$

  $\neq P(A)P(C) = 0.250000499998\ldots$

Output independent of input?

- A minor change in the probabilities of A and B from P(A) = P(B) = 0.5 to P(A) = P(B) = 0.500001 destroyed the independence of A and C!
- It would be hard to distinguish between the two cases via experimentation
- The occurrence of stochastic independence of A and C does not imply that A and C are physically independent
- The output of an XOR gate does depend on its input

Independence of three events I

- Events A, B, and C are said to be mutually independent if all four of the following conditions hold:
  - P(AB) = P(A)P(B)
  - P(AC) = P(A)P(C)
  - P(BC) = P(B)P(C)
  - P(ABC) = P(A)P(B)P(C)
- It would appear that the first three conditions imply the fourth, or the fourth implies the first three, but this is not true

Example (fair tetrahedral die)

- Example: The four triangular faces of a tetrahedral fair die are marked with the numbers 2, 3, 5, and 30 respectively. The outcome is the number on the bottom face when the die is rolled
- A, B, and C are events that the outcome is a multiple of 2, 3, and 5 respectively
- P(A) = P(B) = P(C) = 1/2
- P(AB) = P(AC) = P(BC) = P(ABC) = 1/4
- First three equations hold but not fourth

Example (loaded tetrahedral die)

- Example: Now suppose that the die is loaded such that the four outcomes 2, 3, 5, and 30 have probabilities 11/24, 7/24, 5/24, and 1/24 respectively
- A, B, and C are events that the outcome is a multiple of 2, 3, and 5 respectively
- P(A) = 1/2; P(B) = 1/3; P(C) = 1/4
- P(AB) = P(AC) = P(BC) = P(ABC) = 1/24
- P(ABC) = P(A)P(B)P(C)
- Fourth equation holds but not first three

Independence of three events II

- An alternative definition that generalizes more easily is that A, B, and C are independent if P(ABC) = P(A)P(B)P(C) holds, and each subset of two events is also independent
- Since there are 3 subsets of two events, we get the other three conditions
  - P(AB) = P(A)P(B)
  - P(AC) = P(A)P(C)
  - P(BC) = P(B)P(C)

Independence of three events III

- An third definition that generalizes more easily is that A, B, and C are independent if each of the following 8 equations holds:

  P(A*B*C*) = P(A*)P(B*)P(C*)

  where A* denotes either A or A^c, B* denotes either B or B^c, and C* denotes either C or C^c
- All the three definitions are equivalent: the latter two generalize more easily

Independence of multiple events I

- $\{A_1, A_2, \ldots, A_n\}$ is said to be a collection of mutually independent events if
  - $P(A_1A_2\ldots A_n) = P(A_1)P(A_2)\ldots P(A_n)$
  - every subcollection containing two or more of the A_i's is also a collection of independent events
- This is a recursive definition: the product rule applies to the "big" intersection and also to all smaller ones as well

  $P(A_iA_j\ldots A_m) = P(A_i)P(A_j)\ldots P(A_m)$

Independence of multiple events II

- $\{A_1, A_2, \ldots, A_n\}$ is said to be a collection of mutually independent events if all 2^n of the following equations hold:

  $P\{(A_1)^*(A_2)^* \ldots (A_n)^*\}$

  $= P\{(A_1)^*\}P\{(A_2)^*\} \ldots P\{(A_n)^*\}$
- Here, as before, $(A_i)^*$ represents either A_i or $(A_i)^c$ (the same on both sides of each equation)
- Two choice for each $(A_i)^*$ 2^n equations

Independence of multiple events III

- $P\{(A_1)^*(A_2)^* \ldots (A_n)^*\}$

  $= P\{(A_1)^*\}P\{(A_2)^*\} \ldots P\{(A_n)^*\}$

  implies that the product rule also applies to every subcollection containing two or more of the $(A_i)^*$'s
- $P\{(A_i)^*(A_j)^* \ldots (A_m)^*\}$

  $= P\{(A_i)^*\}P\{(A_j)^*\} \ldots P\{(A_m)^*\}$

Independence of multiple events IV

- We have looked at two different definitions of independent events
- The definitions are equivalent
- Both can be used simultaneously
- The important point to keep in mind is that independence applies to all the subsets and it applies even if we are using the complements of the events

The union of independent events

- It is easy to compute the probability of the intersection of independent events
- What about their union?
- The obvious answer is to use the principle of inclusion/exclusion: Include the probability of the events, exclude the probabilities of the pairwise intersections, include the probabilities of the triplewise intersections, …
- All the needed probabilities can be computed via the product rule!

Easier said than done …

- Use of the principle of inclusion/exclusion requires a lot of computation!
- Use DeMorgan's theorem instead
- $P(A_1 \quad A_2 \quad \ldots \quad A_n)$

  $= 1 - P\{(A_1 \quad A_2 \quad \ldots \quad A_n)^c\}$

  $= 1 - P\{(A_1)^c(A_2)^c \ldots (A_n)^c\}$

  $= 1 - P((A_1)^c)P((A_2)^c) \ldots P((A_n)^c)$

  $= 1 - (1-P(A_1))(1-P(A_2)) \ldots (1-P(A_n))$
- DO NOT expand this last expression!

Boolean functions

- Any Boolean function of $\{A_1, A_2, \ldots, A_n\}$ is independent of any other Boolean function as long as they do not include any events in common
- Example: If A, B, C are independent events, then A B is independent of C
- Example: If {A,B,C,D,E,F,G,H} are independent events, then A C, B H, D, and E FG are independent events

Boolean functions

- Example: If {A,B,C,D,E,F,G,H} are independent events, then A C, B H, D, and E FG are independent events
- P((A C)(B H)D(E FG))

  = P(A C)P(B H)P(D)P(E FG)
- P(A C) = P(A) + P(C) – P(A)P(C)
- P(B H) = P(B) + P(H) – P(B)P(H)
- P(E FG)= P(E) + P(FG) – P(EFG)

  = P(E) + P(F)P(G) – P(E)P(F)P(G)

Boolean functions (continued)

- When events are shared among Boolean functions, independence cannot be guaranteed. However, problem analysis is still possible in conjunction with Karnaugh maps
- Example: If A, B, and C are independent, A C and B C are not independent events
- P((A C)(B C)) = P(C AB) = P(C ABC^c)
  = P(C) + P(ABC^c) = P(C) + P(A)P(B)P(C^c)

Independent Random Variables

- **X** is a random variable defined on the simple experiment. It maps a_5 to **X**(a_5)
- $\mathbf{X}_i$ denotes the number observed on the i-th subexperiment of the compound experiment
- ($\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$,…) is called a random vector
- If the outcome is (a_5, a_2, a_7, a_9, a_1, …), then $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, $\mathbf{X}_4$, $\mathbf{X}_5$,**…** have values **X**(a_5), **X**(a_2), **X**(a_7), **X**(a_9), **X**(a_1),… etc

Independent random variables

- For repeated independent trials, the random variables $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, $\mathbf{X}_4$,… are said to be independent random variables
- If **X** is a discrete random variable, then for repeated independent trials, we have

P($\mathbf{X}_1 = a_5$, $\mathbf{X}_2 = a_2$, $\mathbf{X}_3 = a_7$, $\mathbf{X}_4 = a_9$, …) =
= P($\mathbf{X}_1$=a_5)P($\mathbf{X}_2$=a_2)P($\mathbf{X}_3$=a_7)P($\mathbf{X}_4$=a_9)…

- This is just
  P(A, B, C, A^c, …) = P(A)P(B)P(C)P(A^c)…

X_i on the compound experiment

- The random variable $\mathbf{X}_i$ was defined on the i-th subexperiment
- We can also view $\mathbf{X}_i$ as being defined on the compound experiment
- $\mathbf{X}_i$ maps the outcome

  (✶, ✶, ✶, … , a_2, ✶, ✶, …)

  to the number X(a_2)
- Here ✶ means that we don't care what outcome is in that position

Independence on compound expt

- Random variables $\mathbf{X}_1$, $\mathbf{X}_2$, … viewed as being defined on the compound experiment are independent
- We express their independence via

P($\mathbf{X}_1 = a_5$, $\mathbf{X}_2 = a_2$, $\mathbf{X}_3 = a_7$, $\mathbf{X}_4 = a_9$, …)
= P($\mathbf{X}_1$=a_5)P($\mathbf{X}_2$=a_2)P($\mathbf{X}_3$=a_7)P($\mathbf{X}_4$=a_9)…

- This notion can be applied to other (not necessarily compound) experiments as well

Independent random variables

- Definition: The discrete random variables **X** and **Y** are said to be (stochastically) independent if, for all real numbers a and b,

  P({**X** = a}∩{**Y** = b}) = P{**X** = a}•P{**Y** = b}
- If **X** and **Y** are physically independent (e.g. repeated trials), then the above holds
- But it is possible for the above equation to hold even though the random variables are provably physically dependent

More generally, …

- Definition: The discrete random variables **X**, **Y**, **Z**, … are said to be (stochastically) independent if, for all real numbers a, b, c, …

  P({**X** = a}∩{**Y** = b}∩{**Z** = c}…)

  = P{**X** = a}•P{**Y** = b}•P{**Z** = c}…
- Once again, this equation can hold for random variables that are provably physically dependent

Summary

- We have discussed independence of events
- We have tried to distinguish between physical independence, a property of the events,and stochastic independence, a property of the probability measure
- Mutual independence and mutual exclusion are completely different!
- We have studied the use of independence in probability calculations

