Probability of Compound Events

Shantanu Dutt
ECE Dept.
Uinv. of Illinois at Chicago
Basics—Mutually Exclusive Events

- \( p(A) \) will denote prob. of event \( A \).
- Two events \( A, B \) are mutually exclusive (ME) if one event happening precludes the possibility of the other. In other words \( p(\text{both } A \text{ and } B \text{ happens}) = p(A \cap B) = 0 \).
- E.g., Event \( D_i \) is a stranger’s b’day being on the \( i \)’th day of the week, \( 1 \leq i \leq 7 \). \( p(D_i) = 1/7 \) for each \( i \). However, \textit{given that the event } \( D_4 \) \textit{has happened} (the stranger’s b’day is found to be on Thurs), the prob. of the other events are 0. Thus the \( D_i \)’s are ME.
- Another ex. is blocks of code in an if-then-else chain. The event that a particular block \( B_j \) will be executed in the current pass through the if-then-else chain is ME w/ the execution of other blocks.
- If \( A \) and \( B \) are ME, then the \( p(\text{either } A \text{ or } B \text{ happens}) = p(A \cup B) = p(A) + p(B) \); see Fig. 1.
- If \( A, B \) are not ME, then \( p(A \cup B) = p(A) + p(B) - p(A \cap B) \), since we are counting the event \( p(A \cap B) \) twice in \( p(A) + p(B) \); see Fig. 2.
- Similarly, \( p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(B \cap C) - p(A \cap C) + p(A \cap B \cap C) \), since we included \( p(A \cap B \cap C) \), 3 times in \( p(A) + ... + p(B) \) but also excluded it 3 times in \( -p(A \cap B) - p(B \cap C) - p(A \cap C) \), and so need to add it back once; see Fig. 3.
- The general formulation for \( p(\cup_{i=1}^n E_i) \) is obtained from a generalization of the above –inclusion-exclusion pattern and is formally called the \textit{inclusion-exclusion principle}; see http://en.wikipedia.org/wiki/Inclusion%E2%80%93exclusion_principle

![Fig. 1](image1.png)

![Fig. 2](image2.png)

![Fig. 3](image3.png)
Basics—Mutually Exclusive Events

• Two events A, B are independent if any one of the events happening does not affect the probability of the other. This is expressed as \( P(A/B) = P(A) \), and \( P(B/A) = p(B) \), where \( p(A/B) \) is the probability of event A happening given that B has happened.

• E.g., Event A = it rains today; Event B = there are dark clouds in the sky; Event C = it is very sunny; Event D = today is Wed.
  – Then A and D are independent, i.e., \( P(A) \) is not affected by D happening, or \( p(A/D) = p(A) \). Also, \( p(D/A) = p(D) \).
  – Similarly B and D, and C and D are independent.
  – But A and B are not independent, and statistics will tell us that \( p(A/B) > p(A) \), and \( p(B/A) > p(B) \).
  – Similarly, A and C are not independent, and \( p(A/C) < p(A) \) and \( p(C/A) < p(C) \).

• If A and B are independent, then \( p(A \cap B) = p(A)*p(B) \). Otherwise, \( p(A \cap B) = p(A)*p(B/A) = p(B)*(A/B) \). E.g., in above ex., say, \( p(A) = 0.2 \) and \( p(B) = 0.25 \). Then \( p(A \cap B) \) will be much higher than \( p(A)*p(B) = 0.05 \), since the probability that both A and B happens will be very close to \( p(A) = 0.2 \) (i.e., \( p(A/B) \) is close to 1), since if it is raining it is very likely that the clouds are very dark.

• The determination of whether 2 events are independent are determined logically based on the given even scenario OR statistically.

• Similarly the formulation of \( p(A/B) \) is determined based on the underlying \( p(A) \) and \( p(B) \) formulations OR from statistics.

• Once independence or non-independence and \( p(A/B) \) are determined by non-probabilistic methods, \( p(A \cap B) \) can be determined.

Fig. 4: Note that if \( A \cap B \) is not empty it does not have any bearing on whether A and B are indep. or not, which are determined by other means.
Exclusive Atomic Events

- **Scenario**: There is a box with a red ball in it. A group of \( n \) randomly picked individuals are picked and placed equidistant from the box. At the count of 3, they are to rush to the box and take the ball out. Clearly, only one person can get the ball. See Figs. 5 and 6.

- The atomic event \( E_i \) is person \( i \), \( 1 \leq i \leq n \). \( p(E_i) = \frac{1}{n} \).

- Hence the prob. that either person \( i \) or person \( j \) gets the ball \( p(E_i \cup E_j) = p(E_i) + p(E_j) = \frac{2}{n} \). Similarly, the prob. that any person in a subset of \( k \) people gets the ball = \( \frac{k}{n} \).

- Consider event \((E_i)' = \) event that \( i \) does not get the ball. \( p((E_i)') = 1 - p(E_i) \) \([\text{by definition of the complement event}] = \frac{n-1}{n} \).

- \((E_i)' \) and \((E_j)'\) are not ME (both can happen), and they are also not independent. The latter is so, since if \( E_i \) does not get the ball, then the prob. of \( E_j \) increases to \( 1/(n-1) \), i.e. \( p(E_j/(E_i)') = 1/(n-1) \neq p(E_j) \). Thus \( p((E_i)'/(E_j)') = 1 - p(E_j/(E_i)') = 1 - (1/(n-1)) = (n-2)/(n-1) \neq p((E_i)') \).

- Thus \( p((E_i)' \cap (E_j)') = p((E_j)')p((E_i)'/(E_j)') = ((n-1)/n)*((n-2)/(n-1)) = (n-2)/n \)

- Prob. that one of person \( i \) or \( j \) gets the ball = \( p(E_i \cup E_j) \) can also be derived as (whether these events are ME or not): \( 1 - \text{prob.}(\text{none of them get the ball}) = 1 - p((E_i)' \cap (E_j)') = 1 - (n-2)/n = 2/n \)

- If 2 events \( A, B \) are not ME then \( p(A \cup B) = 1 - p(A' \cap B') \) is an easier derivation than determining \( p(A \cup B) \) using the inclusion-exclusion principle, especially if the number of events in the union is > 2.

- For \( A, B, C \), \( p(A \cup B \cup C) = 1 - p(A' \cap B' \cap C') = 1 - p(A'/ B' \cap C') *p(B' \cap C') \)

- For the above scenario, this is \( 1 - (1-p(A/ B' \cap C'))*(n-2)/n = 1-(1-1/(n-2)*(n-3))/n = 1 - (n-3)/n = 3/n \)

- Thus for the ex. in Fig. 5, \( p(E_1 \cup E_3) \) can be obtained as either: \( p(E_1) + p(E_3) \) (as they are ME) = \( 2/3 \) or the more general (since ME atomic events are rare), \( 1 - p(E_1' \cap E_3') = 1 - p(E_1'/E_3')p(E_3') = 1 - (1-p(E_1/E_3'))(1-p(E_3)) = 1-(1-1/2)(1-1/3) = 1-(1/2)*(2/3) = 2/3 \)
Exclusive Atomic Events (cont’d)

• Scenario: In an indirect tree topology, the top switch (see Fig. 7) is the sole communication route between the \((P/2)\)-processor subsets on its left and right subtree. We need to know the message load on this switch, when each processor sends a msg. to another random processor. We do this by focusing on a single processor \(v\) (=4 in Fig. 7) and determining what the prob. is of its msg. going to the \((P/2)\)-proc. subset on the other side of the tree.

• The atomic events are \(\{Mi: \text{proc. }i \text{ gets the msg. from } v\}\), w/ only \(M_v = \phi\), since \(v\) does not send a msg. to itself. For other \(Mi\)'s, \(q = p(M_i) = 1/p-1\) (=1/5 in Fig. 7), as the msg. is sent to a random dest. & thus the prob. distr. is uniform.

• The events \(Mi\) are also clearly ME, as there is only 1 msg. and only 1 proc. will get it.

• Thus \(p(M_1 U M_2) = p(M_1)+p(M_2) = 2/(p-1)\) (=2/5 in ex.). In general, \(p(a \text{ proc. in the other } (P/2)-\text{proc. subset getting the msg. from } v) = p(M_1 U M_2 \ldots U M_{P/2}) = p(M_1)+p(M_2)+\ldots+p(M_{P/2}) = (p/2)*q\) (= 3/5 for the Fig. 7 ex.)

• Similarly, \(p(a \text{ proc. in the other } (P/4)-\text{subset within } v's \,(P/2)-\text{subset getting the msg.}) = (p/4)*q\).

• Also, as in the prev. “ball-grabbing” scenario, \(Mi\)'s are not independent, buy similarly, \(p(M_1 U M_2 \ldots U M_{P/2})\) can also be derived as \(1 – p(M'_1 \cap M'_2 \cap \ldots \cap M'_{P/2}) = 1 – (p(M'_1)*p(M'_2)/(M'_1)*\ldots*p(M'_P/(M'_1 \cap M'_2 \ldots M'_{(P-1)/2}))) = 1 – [(p-2)/(p-1)]*[((p-3)/(p-2)]*\ldots*[(p-p/2 -1)/(p-p/2)] = (p/2)/\ldots/(p-1) = (p/2)*q\).

• Thus the msg load (or prob.) on the top switch in a random msg. pattern is \(p^*(p/2)*q = p^2q/2\). On the switch below the root/top one, msg load is \((p/2)*(p/4)*q = p^2q/8\). Thus msg load on the top switch is 4 times that of the switch below it, and so on, leading to the rationale for a fat-tree topology in which the switch size and \(\#\) of links doubles (or increases by a factor > 1 and <= 2) so that msg. latency is not high due to contention/collision. It doubles instead of increasing by a factor of 4, as msg. patterns are mostly not random but more structure and more localized (e.g., in recursive reduction—what is the relative load here?)
Non-ME Atomic Events

- Scenario: In an indirect tree topology, the top switch (see Fig. 9) is the sole communication route between the \((P/2)\)-processor subsets on its left and right subtree. We need to know the message load on this switch, when each processor sends a msg. to another random processor. We consider the left \((P/2)\)-subset sending msg. randomly and determine the prob. of at least 1 msg reaching the right \((P/2)\)-subset. This will be representative of the load on the top sw.

- Assume we did not do the previous analysis for a single-proc. msg.

- The atomic events are \(\{M_{i,j}: \text{proc. }i \text{ gets the msg. from }j\}\), w/ only \(M_{j,j} = \phi\), for each \(j\), since a proc. does not send a msg. to itself. For other \(M_{i,j}\)'s, \(p(M_{i,j}) = 1/p-1 = q (=1/5 \text{ in Fig. 9})\), as the msg. is sent to a random dest. & thus the prob. distr. is uniform.

- The events \(M_{i,j}\) for a constant \(j\) are clearly ME, as there is only 1 msg. from \(j\) and only 1 proc. will get it.

- However, all \(M_{i,j}\)'s (i.e., when both \(i\) & \(j\) are variables) are not ME’s as more than one (in fact up to \(P/2\)) of these events can happen simultaneously.

- Thus \(p(\text{at least 1 msg. coming to left \((P/2)\)-subset from right \((P/2)\)-subset}) = p(U_{i \text{ in left}, j \text{ in right } M_{i,j}}) = \sum_{i \text{ in left}, j \text{ in right}} p(M_{i,j})\). Further, using the union/sum approach w/ the inclusion-exclusion principle becomes very cumbersome (though it will yield the correct result).

- However, using the “complement of the intersection of the complements” approach, \(p(U_{i \text{ in left}, j \text{ in right } M_{i,j}}) = 1 - p(\bigcap_{i \text{ in left}, j \text{ in right }}(M_{i,j})') = 1 - p(\bigcap_{i \text{ in left}}(\bigcap_j \text{ in right } (M_{i,j})'))\).

- Let \(M_i\) be the event that \(i\) gets at least 1 msg from the left. Thus \(p((M_i)' = p(i \text{ gets no msgs from the left}) = p(\bigcap_j \text{ in right } (M_{i,j})'))\).

- Event \((M_i,j)' = i\) does not get a msg. from \(j\). Any two events \((M_i,j)'\) and \((M_i,k)'\), \(j \neq k\), are independent as the occurrence of one does not affect that of the other. Thus \(p((M_i)' = p(\bigcap_j \text{ in right } (M_{i,j})') = \prod_j \text{ in right } p((M_{i,j})') = \prod_j \text{ in right } (1 - (1/(P-1)) = \prod_j \text{ in right } (P-2)/(P-1)

\[= [(P-2)/(P-1)]^{P/2}\]
Non-ME Atomic Events (contd.)

- Event \( (\text{Mi},j)' \) = \( i \) does not get a msg. from \( j \). Any two events \( (\text{Mi},j)' \) and \( (\text{Mi},k)' \), \( j \neq k \), are independent as the occurrence of one does not affect that of the other. Thus \( p((\text{Mi})') = p(\bigcap_{j \text{ in right}} (\text{Mi},j)') = \Pi_{j \text{ in right}} p((\text{Mi},j)') = \Pi_{j \text{ in right}} (1 - (1/(P-1))) = \Pi_{j \text{ in right}} (P-2)/(P-1) = [(P-2)/(P-1)]^{p/2} \)

- However, \( (\text{Mi})' \) and \( (\text{Mk})' \), \( i \neq k \), are not independent: no msgs. recvd. at \( i \) (i.e., \( (\text{Mi})' \) occurring) increases \( \text{Mk} \), since it increases each component \( \text{Mk},j \) to \( (1/(P-2)) \), and thus decreases \( (\text{Mk})' \).

- Thus \( p((\text{Mi})' \cap \text{Mk}') = p((\text{Mi})')p((\text{Mk}')/\text{Mi}') \)

- \( p((\text{Mk}')/\text{Mi}') = p(\bigcap_{j \text{ in right}} ((\text{Mk},j)')/\text{Mi}') \). Each \( p((\text{Mk},j)'/\text{Mi}') \) is indep. across changing \( j \)’s.

- Thus \( p((\bigcap_{j \text{ in right}} ((\text{Mk},j)')/\text{Mi}')) = \Pi_{j \text{ in right}} p((\text{Mk},j)')/\text{Mi}') = [(P-2)/(P-1)]^{p/2} \)

- Extrapolating in an obvious way, \( p(\bigcap_{l \text{ in left}} ((\bigcap_{j \text{ in right}} ((\text{Mi},j)'))) = p(\bigcap_{l \text{ in left}} \text{Ml}') = [(P-1 - (P/2))/(P-1)]^{p/2} = [(P/2)-1]/(P-1)]^{p/2} \)

- \( p(U_{l \text{ in left, j in right}} \text{Mi},j) = 1 - p(\bigcap_{l \text{ in left, j in right}} ((\text{Mi},j)')) = 1 - p(\bigcap_{l \text{ in left}} ((\bigcap_{j \text{ in right}} ((\text{Mi},j)'))) = 1 - [(P/2)-1]/(P-1)]^{p/2} \)

- For the ex. in Fig. 9 this prob. = \( 1 - (2/5)^3 = (125 - 8)/125 = 117/125 = 0.936 \).

- For Fig. 10 it is \( 1 - (3/7)^4 = (2401 - 81)/2401 = 2320/2401 = 0.966 \).

- Similarly, \( p(\text{a proc. in one of the (P/4)-subset within the left (P/2)-subset getting at least msg. from its other (P/4)-subset}) \) [this will be via its 2nd topmost switch = \( 1 - [(3P/4)-1]/(P-1)]^{P/4} \) = (for Fig. 10) \( 1 - (5/7)^2 = 24/49 = 0.49 \) (almost \( 1/2 \) of that for the top switch). Thus again, load (in terms of prob. of \( \geq 1 \) msg. through it) on the top switch is about \( 2X \) that of switch below it and so forth.

Fig. 10: Scenario in which processors in the left (P/2)-proc. subset sends one message each to a random destination. Q: What is the prob. of at least msg. going through the top switch? This prob. is representative of the msg. load on the top switch under this scenario.
Non-ME Atomic Events (contd.)

- Via union and incl-excl principle $p(\text{at least 1 msg. to left from right for Fig. 9} ) = (\text{see Fig. 11}) p(M1)+p(M2)+p(M3) - p(M1 \cap M2) - p(M1 \cap M3) - p(M2 \cap M3) + p(M1 \cap M2 \cap M3)$. This will be very involved:

- Each component is quite involved, e.g., $p(M1) = 1 - p(M1')$ (which itself is using the compl. of compl. approach) = 1 - $\left(\frac{P-2}{P-1}\right)^{P/2}$ = (for Fig. 9) = 1 - $(4/5)^3 = (125 - 64)/125 = 61/125 = p(M2) = p(M3)$

- $p(M1 \cap M2) = p(M1) - p(M1 \cap M2')$; $p(M1 \cap M2') = p(M2') - p(M1' \cap M2')$

- From earlier, $p(M1' \cap M2') = \left(\frac{P-3}{P-1}\right)^{P/2} = (3/5)^3 = 27/125$

- Thus $p(M1 \cap M2) = p(M1) - p(M1 \cap M2') = 61/125 - 27/125 = 34/125$

- $p(M1 \cap M2' \cap M3') = p(M2' \cap M3') * p(M1/(M2' \cap M3')) = (27/125)*(19/27)$

- Fig. 9: Scenario in which processors in the left (P/2)-proc. subset sends one message each to a random destination. Q: What is the prob. of at least msg. going through the top switch? This prob. is representative of the msg. load on the top switch under this scenario.

- $p(M1 \cup M2 \cup M3) = p(M1)+p(M2)+p(M3) - p(M1 \cap M2) - p(M1 \cap M3) - p(M2 \cap M3) + p(M1 \cap M2 \cap M3) = 3*61/125 + 0.048 = 0.936$ (same as in prev. slide)