Review.

1. Mean, Variance, Moments

\[ \text{mean, average, expected value}: \ E[X] = \sum_i u_i P_X(u_i) \]

\[ \text{LOTUS: } E \left[ g(X) \right] = \sum_i g(u_i) P_X(u_i) = \mu_X \]

2. Variance & standard deviation:

\[ \text{Variance: } \text{Var}(X) = E \left[ (X-\mu_X)^2 \right] = \sum_i (u_i-\mu_X)^2 P_X(u_i) \]

\[ = E[X^2] - \mu_X^2 \]

\[ \text{Standard deviation: } \sigma_X = \sqrt{\text{Var}(X)} \]

\[ \text{Standardized version: } Y = \frac{X-\mu_X}{\sigma_X} \]

\[ \text{has no unit, } E[Y] = 0, \text{Var}(Y) = 1 \]

3. Properties: follows by LOTUS

\[ \text{Var}(aX) = a^2 \sigma_X^2, \quad \sigma_{aX} = |a| \sigma_X \]

\[ E[aX] = a \mu_X, \quad E[X+b] = \mu_X + b \]

\[ \text{Var}(X+b) = \text{Var}(X) \]

4. \(i\)'th moment = \(E[X^i]\), \(i \geq 1\) integer

5. Conditional probability:

\[ P(B|A) = \begin{cases} \frac{P(AB)}{P(A)}, & P(A) > 0 \\ \text{undefined}, & P(A) = 0 \end{cases} \]

Given the information that \(A\) happened, how likely is \(B\)?

Frequency interpretation

We ran an experiment \(N\) times.

\[ N_A = \# \text{ of times } A \text{ happened} \]
\[ N_A = \# \text{ of times } A \text{ happened} \]
\[ N_{AB} = \# \text{ of times } AB \text{ happened} \]
\[ P(B|A) \propto \frac{N_{AB}}{N_A} \]

- \( P(B|A) \) is a probability measure, satisfies axioms of probability. Given \( P(A) > 0 \),
  1. \( P(\Omega|A) = 1 \)
  2. \( P(B|A) > 0 \) for all \( B \in F \)
  3. If \( B_1, B_2, \ldots \) are mutually exclusive \( P(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i) \)

**Today:**

- Bayes rule
  1. Independence

Bayes rule: \( P(B|A) \cdot P(A) = P(A|B) \cdot P(B) = P(AB) \) given \( P(A), P(B) > 0 \)
  in particular \( P(B|A) = \frac{P(AB)P(B)}{P(A)} \)

3. Independence:

  **Def:** \( A \) & \( B \) are independent if \( P(AB) = P(A)P(B) \).

  Intuition: suppose that \( P(A) > 0 \). Then \( A \) & \( B \) are independent if knowledge of \( A \) has no implication on \( B \), i.e.,
  \[ P(B|A) = P(B) \]

  note: \( P(AB) = P(A)P(B) \) is just an equation:
  * roll a die, \( A = \{2, 4, 6\} \), \( B = \{5, 6\} \), \( AB = \{6\} \)
A and B are independent: \( P(AB) = P(A) \cdot P(B) = \frac{1}{6} = P(A) \cdot P(B) \)

**Def.** A, B, C are pairwise independent if

\[
P(AB) = P(A) \cdot P(B) \]
\[
P(AC) = P(A) \cdot P(C) \]
\[
P(BC) = P(B) \cdot P(C) \]

**Def.** A, B, C are independent if they are pairwise independent and

\[
P(ABC) = P(A) \cdot P(B) \cdot P(C) \]

**Def.** \( A_1, A_2, \ldots, A_n \) are independent if

\[
P(A_{i_1} A_{i_2} \ldots A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \ldots P(A_{i_k})
\]

for any \( 2 \leq k \leq n \) and \( 1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq n \)

**note:** A, B, C being pairwise independent won't imply independence.

*roll a 9-sided die. sides are \{1, 2, \ldots, 9\}. outcomes are equilikely.*

\[
A = \{1, 2, 3\} \quad B = \{3, 4, 5\} \quad C = \{2, 4, 6\}
\]

\[
P(A) = P(B) = P(C) = \frac{3}{9}
\]

\[
P(AB) = P(BC) = P(AC) = \frac{1}{9}
\]

\[
P(ABC) = 0
\]

**Implications:**

1. If \( A \) is independent of \( B \), then \( A^c \) is independent of \( B \).

\[
P(B) = P(B \cap (A \cup A^c)) = P(B \cap A) + P(B \cap A^c)
\]

\[
\rightarrow P(A^c \cap B) = (1 - P(A)) \cdot P(B) = P(A^c) \cdot P(B)
\]

2. If \( A, B, C \) are independent, then \( A \) is independent of \( B \cup C \)

\[
P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C))
\]
\[ P(AB) + P(AC) - P(ABC) \]
\[ P(A)(P(B) + P(C) - P(BC)) \]
\[ P(A)P(B\cap C) \]

In particular, A is independent of any combination of B & C.

Above properties generalizes to \( A_1, A_2, \ldots, A_n \).

\( B_1 \) : combination of \( n_1 \) out of \( \{A_1, \ldots, A_n\} \)

\( B_2 \) : combination of \( n_2 \) out of \( \{A_1, \ldots, A_n\} \) that did not appear in \( B_1 \).

\[ \vdots \]

\( B_k \) : combinations of \( n_k \) out of \( \{A_1, \ldots, A_n\} \) that did not appear in \( B_1, \ldots, B_{k-1} \).

Then \( B_1, B_2, \ldots, B_k \) are independent.

---

Important remarks:

1. Independence vs mutually exclusive:
   - We say A & B are independent, if \( P(AB) = P(A)P(B) \).
   - We say A & B are mutually exclusive, if \( A \cap B = \emptyset \).

In particular, being independent relies on the underlying probability measure \( P \); being mutually exclusive has nothing to do with probability measure and it is a set level definition.

2. Intersection vs conditional probability:
   - \( A \cap B \) : the event \( A \text{ AND} B \). It is a set level definition.
   - \( P(B|A) \) : the probability of event B \( \text{ GIVEN} \) that A has happened. It depends on a probability measure \( P \).
Here is a semi-example:

Suppose that Jane lost his dog, and the dog is in forest A or forest B. Jane uses a coin to decide which forest to search.

- What is the probability that Jane will search forest A & find the dog?

  the question is asking for

  \[ P(\{\text{Jane will search forest A} \} \cap \{\text{dog is in A and search is successful} \}) \]

- After tossing the coin, Jane decides to search forest A. What is the probability that Jane will find the dog?

  the question is asking for

  \[ P(\{\text{dog is in A and search is successful} \} \mid \{\text{Jane will search forest A} \}) \]