Lecture 4,
August 30

Conditional Probability and independence

Conditional probability for computing the probability of intersections

It follows from our definition of conditional probability that

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

Example: Urn 1 has 3 white & 1 red balls; Urn 2 has 2 white & 2 red balls. Select 1 ball randomly from Urn 1 & place it into Urn 2. Then, select 1 ball randomly from Urn 2. What's the probability that ball selected from Urn 2 is red?

3W
1R
Urn 1

2W
2R
Urn 2

$$P(\text{2nd } R) = P(\text{1st Red \& 2nd Red} \cup \text{ OR 1st W \& 2nd Red})$$

$$= P(\text{1st Red \& 2nd Red}) + P(\text{1st W \& 2nd Red})$$

$$= P(\text{1st Red})P(\text{2nd Red} \mid \text{1st Red}) + P(\text{1st W})P(\text{2nd Red} \mid \text{1st W})$$

$$= \frac{1}{4}\,\frac{3}{5} + \frac{3}{4}\,\frac{2}{5}$$

More generally, for events $A_1, A_2, \ldots, A_n$

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1})$$

Example: Suppose we pick 4 cards randomly from a deck of 52 cards having 4 aces

- let event A = we pick 4 aces
- let event A_i = ith pick is an ace ($i = 1, 2, 3, 4$)
- Using an equally likely model approach:

$$P(A) = \frac{1}{\binom{52}{4}} = \frac{4!}{(52)(51)(50)(49)}$$

- Using conditional probabilities

$$P(A) = P(A_1 \cap A_2 \cap A_3 \cap A_4) = \underbrace{P(A_1)}_{\frac{4}{52}}\underbrace{P(A_2|A_1)}_{\frac{3}{51}}\underbrace{P(A_3|A_1 \cap A_2)}_{\frac{2}{50}}\underbrace{P(A_4|A_1 \cap A_2 \cap A_3)}_{\frac{1}{49}}$$

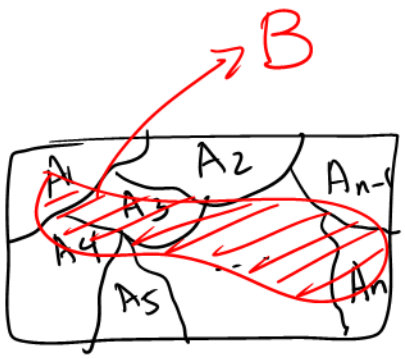
Conditional Probability and independence

Bayes' rule

It is possible to reverse the conditioning of A and B to obtain Bayes' rule:

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

More generally, $A_1, A_2, \ldots,$ is a partition of the sample space S, then we get a general version of Bayes' rule:



$$\underbrace{P(A_i|B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}$$

$$\overset{\text{def}}{=} \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)\,P(B|A_i)}{\sum_{j=1}^{\infty} P(A_j)\,P(B|A_j)}$$

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Example: drug testing at the Olympics

- randomly choose an athlete for testing
- let event U = athlete is a drug user
- let event A = test is positive
- suppose we know $\boxed{P(A|U) = .95}$ and $P(A^c|U^c) = .99$
- if $\boxed{P(U) = 0.005}$ then

$$P(A) = P(\underbrace{\text{test is Positive}}_{A})$$
$$= P(A \cap U) + P(A \cap U^c)$$
$$= P(U)P(A|U) + P(U^c)P(A|U^c)$$

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

$$\begin{aligned} P(U|A) &= \frac{P(A|U)P(U)}{P(A|U)P(U) + P(A|U^c)P(U^c)} \\ &= \frac{(.95)(.005)}{(.95)(.005) + (0.01)(.995)} \\ &\approx .323 \end{aligned}$$

Conditional Probability and independence

Independence

If $P(A|B) = P(A)$ then the occurrence of B doesn't affect the probability of A

It then follows that $P(A \cap B) = P(A)P(B)$ and $P(B|A) = P(B)$

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

$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$

We *define* two events A and B as **independent** if

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

More than two events: $A_1, \ldots, A_n$ are **independent** *if and only if*, for any sub-collection $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ of distinct indices (any $2 \le k \le n$), it holds that

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

- $A_1, \ldots, A_n$ independent $\Rightarrow P(A_i \cap A_j) = P(A_i)P(A_j)$ holds for any $i \neq j$

  But, $P(A_i \cap A_j) = P(A_i)P(A_j)$ for $i \neq j$ *doesn't* imply independent $A_1, \ldots, A_n$

  Example: roll 2 fair dice $\rightarrow 1/6$

  (4,2), (4,6), (4,4)
  (2,2), (2,4), (2,6)
  (4,2), (6,6), (4,4), (6,4)

  Define events A =1st roll is 4; B =2nd roll is 2; C =sum is even;

  $P(A) = \frac{1}{6} \quad P(B) = \frac{1}{6} \quad P(C) = \frac{1}{2}$

  $P(A \cap B) = \frac{1}{36} = P(A)P(B)$

  $\frac{3}{36} = P(A \cap C) = \frac{1}{12} = P(A)P(C)$ $\quad 1/6 \quad 1/2$

  $P(B \cap C) = \frac{1}{12} = P(B)P(C)$

  $P(A \cap B \cap C) \neq P(A)P(B)P(C)$

A, B, C are pairwise independent, BUT A, B, C are NOT independent.

- $A_1, \ldots, A_n$ independent $\Rightarrow P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n)$

  But, $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n)$ holding *doesn't* imply independent $A_1, \ldots, A_n$

  Example: roll 2 fair dice

  Define events $A =$"double"; $B =$sum between 7 & 11; $C =$sum is 2,7 or 8

$$
\begin{aligned}
P(A) &= \frac{1}{6} \qquad P(B) = \frac{1}{2} \qquad P(C) = \frac{1}{3} \\
P(A \cap B \cap C) &= P(\text{roll } (4,4)) = \frac{1}{36} = P(A)P(B)P(C) \\
P(A \cap B) &= P(\text{roll } (4,4) \text{ or } (5,5)) \neq P(A)P(B)
\end{aligned}
$$

Conditional Probability and independence

Independence example

The assumption of independence of events allows the computation of joint occurrences of events through simple calculations

Example: "Parallel System Reliability"

Suppose one can send a message through any one of 3 independent communications channels. The communications system is "up" if at least one channel is "up."

Suppose that at any time

$P(\text{channel } A \text{ is up}) = 0.99, \quad P(\text{channel } B \text{ is up}) = 0.98, \quad P(\text{channel } C \text{ is up}) = 0.97.$

Find $P(\text{system is up})$.

$$\begin{aligned} P(\text{System is Up}) &= 1 - P(\text{system down}) \\ &= 1 - P(A^c \cap B^c \cap C^c) \\ &= 1 - P(A^c)P(B^c)P(C^c) \\ &= 1 - \{(.01)(.02)(.03)\} \\ &= .999994 \end{aligned}$$

Note: If A and B are independent, then $\underbrace{(A^c \text{ and } B^c)}$, $\underbrace{(A \text{ and } B^c)}$, $(A^c \text{ and } B^c)$ are independent

Random variables

Definition

- Thus far, probability on a general sample space S with abstract definition of events

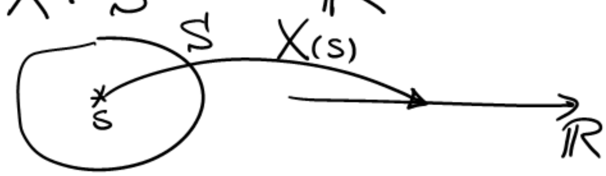
  "transfer sample space S to numbers"

- Statisticians usually are interested in quantitative summaries

- *Definition:* A **random variable** (r.v.) X is a function defined on a sample space S that associates a real number with each outcome in S

  That is, for each $s \in S$, we have $X(s) \in \mathbb{R}$

  $X : S \longrightarrow \mathbb{R}$

  In function notation: $X : S \to \mathbb{R}$

- We usually suppress the dependence of X on $s \in S$ and write $X = X(s)$

Random variables

Examples $X: S \longrightarrow \mathbb{R}$

- Toss three coins

$$
\begin{aligned}
S \equiv \{ & s_1 = \underbrace{HHH}, s_2 = HHT, s_3 = HTH, s_4 = HTT, \\
& s_5 = THH, s_6 = THT, s_7 = \underbrace{TTH}, s_8 = \underbrace{TTT}\}
\end{aligned}
$$

If $X =$"number of heads," then $X(s_1) = 3$, $X(s_8)=0$

If $Y =$"number of tails before first head," then $Y(s_1) = 0$, $Y(s_7)=2$, $Y(s_8)=3$

- Suppose $\underbrace{S = (0, 1)}$ is the sample space (i.e., generate number between 0 & 1)

$Y = Y(s)$="2nd digit (after decimal point) in a decimal expansion of $s \in S$"

$Y(0.312) = 1$, $Y(0.50)=0$

$Y(0.4999\text{---})=9$

$0! = 1$

$0.xyz\text{----}$

def $Y: \#$