This week: 1.5 Independence
1.7 Some exercises
2.1 Random variables (r.v.) (3.1 binomial r.v.)

Conditional independence

Def. Given C, A and B are conditionally indep. if $P(AB|C) = P(A|C)\,P(B|C)$.

Ex 1. Rare disease affects 1 out of 1000 people. A test for this disease is positive for 2% of healthy and 99% of ill people.
J is tested twice and both tests are positive. Find probability that J is healthy.

Answer. H = "J is healthy", H^c = "ill",
A_i = "ith test positive", $i = 1, 2$.
Assume given H, A_1 and A_2 are conditionally independent. The same given H^c:

$$P(A_1 A_2 | H) = P(A_1|H)\,P(A_2|H) = 0.02^2$$
$$P(A_1 A_2 | H^c) = 0.99^2.$$

Then $$P(H|A_1 A_2) = \frac{P(A_1 A_2 | H)\,P(H)}{P(A_1 A_2|H)\,P(H) + P(A_1 A_2|H^c)\,P(H^c)} =$$

$$= \frac{0.02^2 \; 0.999}{0.02^2 \cdot 0.999 + 0.99^2 \; 0.001} = 0.29$$

<u>Comment.</u> $$P(H \mid A_1 \dots A_n) = \frac{0.02^n \cdot 0.999}{0.02^n \cdot 0.999 + 0.99^n \; 0.001}$$

$$= \frac{999}{999 + \left(\frac{99}{2}\right)^n} \xrightarrow{n \to \infty} 0.$$

1.7 <u>Some examples</u>

<u>First step analysis method</u>

<u>Ex 1.</u> (Gambler's ruin) G starts with k dollars, $0 < k < N$. Fair coin is tossed repeatedly. G wins \$1 if H, and loses \$1 if T. G stops in two cases:

1. No money (G is ruined, 0 is reached)
2. N is reached.

Find probability of the ruin.

<u>Answer.</u> A_k = "G starts with \$k and is ruined".
H_1 = "1st toss is H", $T_1 = H_1^c$ = "1st toss is T".
We want to find $p_k = P(A_k)$.
Total probability law:

$$p_k = P(A_k) = P(A_k \mid H_1)\, P(H_1) + P(A_k \mid T_1)\, P(T_1)$$
$$= p_{k+1} \cdot \frac{1}{2} + p_{k-1} \cdot \frac{1}{2} \quad k = 1, \dots, N-1.$$

$$\begin{cases} p_k = P(\ldots) = \ldots \\ \qquad = p_{k+1} \cdot \frac{1}{2} + p_{k-1} \cdot \frac{1}{2}, \quad k = 1, \ldots, N-1. \\ p_0 = 1, \quad p_N = 0. \end{cases}$$

We get a system of eqns for p_k:

$$(1) \begin{cases} p_k = \frac{1}{2} p_{k+1} + \frac{1}{2} p_{k-1}, \quad k = 1, \ldots, N-1 \\ p_0 = 1, \quad p_N = 0 \end{cases}$$

Solving (1): $2p_k = p_k + p_k = p_{k+1} + p_{k-1}$ implies

$$\boxed{\begin{gathered} b := p_{k+1} - p_k = p_k - p_{k-1}, \quad k = 1, \ldots, N-1 \\ b \text{ to be found} \end{gathered}}$$

$p_2 - p_1 = p_1 - p_0$

$p_3 - p_2 = p_2 - p_1$

Note: $p_k = b + p_{k-1} = b + b + p_{k-2} = 2b + p_{k-2} = \ldots = kb + p_0 = kb + 1$

Finding b: $0 = p_N = Nb + 1, \quad b = -\frac{1}{N}$

Answer. $p_k = 1 - \frac{k}{N}, \quad k = 0, 1, 2, \ldots, N$

For instance, $k = 1, N = 100: \quad p_1 = 1 - \frac{1}{100} = 0.99$

$k = 99, N = 100, \quad p_{99} = 1 - \frac{99}{100} = 0.01.$

Remark 1. Note

a) abstract setting helpful

b) conditioning with respect to "first event" could be useful.

Ex 2. Two dice are rolled repeatedly. We are interested in sum of scores.

Find P(5 before 7).

Answer. A = "5 before 7", B_1 = "1st roll 5",
B_2 = "1st roll 7", B_3 = "1st roll neither 5 nor 7."

Then $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3) \times$
$\times P(B_3) = P(B_1) + P(A)\cdot P(B_3)$ ~ eqn. for $P(A)$.

$B_1 = \{(1,4), (2,3), (3,2), (4,1)\}$, $\# B_1 = 4$, $P(B_1) = \frac{4}{36} = \frac{1}{9}$

Similarly $P(B_2) = \frac{6}{36} = \frac{1}{6}$, $P(B_3) = 1 - \left(\frac{1}{9} + \frac{1}{6}\right) = \frac{13}{18}$.

Eqn. for $P(A)$:

$$P(A) = \frac{1}{9} + \frac{13}{18} P(A), \quad P(A) = \frac{18}{5}\cdot\frac{1}{9} = \frac{2}{5} = 0.4.$$

Theoretical exercises

Let $(\Omega, \mathcal{F}, P)$ be probability space.

Main axiom for P is σ-additivity:

If $A_1, A_2, \ldots$ are disjoint, then
$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Continuity property for P. Consider $A_1, A_2 \ldots \in \mathcal{F}$.

a) If $A_n \subset A_{n+1}$ for all n, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

b) If $A_n \supset A_{n+1}$ for all n, then

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n).$$

Why a)?

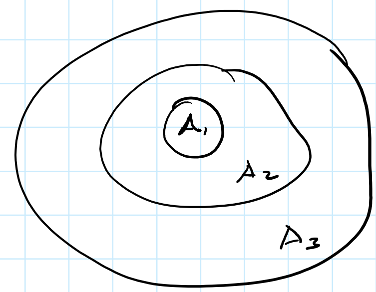
disjoint

$$A_n = \bigcup_{i=1}^{n} (A_i \setminus A_{i-1}), \quad A_0 = \emptyset$$

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{i=1}^{\infty} (A_i \setminus A_{i-1})$$

disjoint

$n = 3$



$$A_3 = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2)$$

So

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} P(A_i \setminus A_{i-1}) = \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i \setminus A_{i-1}) =$$

$$= \lim_{n\to\infty} P(A_n)$$

Why b)? If decreasing, $A_n \supset A_{n+1}$, then $A_n^c \subset A_{n+1}^c$ for all n, and

$$P\left(\bigcap_n A_n\right) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^c\right) = 1 - \lim_{n\to\infty} P(A_n^c) =$$

$= \lim_n [1 - P(A_n^c)] = \lim_n P(A_n).$

<u>Consequences of continuity.</u> Consider $A_1, A_2, \ldots$

Then (i) $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(\bigcup_{i=1}^{n} A_i)$

(ii) $P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(\bigcap_{i=1}^{n} A_i)$

<u>Why?</u> $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty}\left(\bigcup_{i=1}^{n} A_i\right)$ and $\bigcup_{i=1}^{n} A_i$ is increasing sequence: $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(\bigcup_{i=1}^{n} A_i)$

<u>Ex 3.</u> Show that $P(\bigcup_{n=1}^{\infty} A_n) \le \sum_{n=1}^{\infty} P(A_n)$

<u>Answer.</u> $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(\bigcup_{i=1}^{n} A_i) \le$

$\le \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{\infty} P(A_i).$