Elements of Probability	Summer 2019
Independence, conditional probability — 14.03, 2019	
Prof. dr hab. Jacek Cichoń	*Scribe: JK*

1 Overview

The previous lecture included discussion about probability spaces and geometric probability.

This lecture includes basic concept of independence of random variables and conditional probability.

Fact: Suppose that $\Omega = \bigcup_{k=0}^{n} A_k$ is such that: $(\forall i, j)(i \neq j \longrightarrow A_i \cap A_j = \emptyset)$ and $\forall_i (A_i \in \mathcal{S})$. Let $B \in \mathcal{S}$. Then:

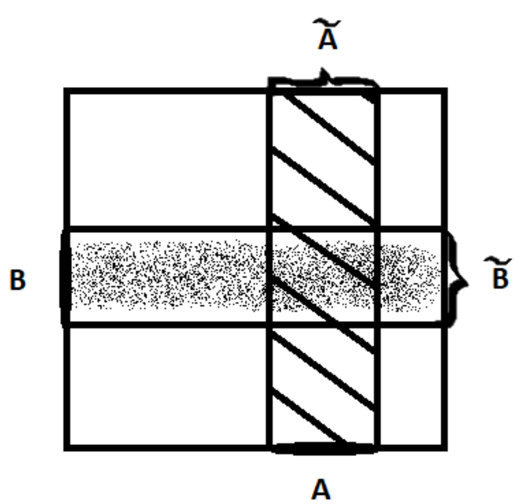
$$P(B) = \sum_{k=0}^{n} P(B \cap A_i) \quad \leftarrow \text{total probability formula}$$

Proof:

$$B = \bigcup_{k=0}^{n} (B \cap A_k) : i \neq j \longrightarrow (B \cap A_i) \cap (B \cap A_j) = \emptyset$$

2 Independence

Definition 1 (independence of random variables)**.** *Random variables $A, B \in \mathcal{S}$ are independent if $P(A \cap B) = P(A)P(B)$.*



Example: Let $\Omega = [0,1]^2$, $A, B \subseteq [0,1]$, $\widetilde{A} = A \times [0,1]$ and $\widetilde{B} = [0,1] \times B$.

Then we have:

$$\lambda^{(2)}(\widetilde{A}) = \lambda^{(1)}(A) \cdot 1 = \lambda^{(1)}(A);$$

$$\lambda^{(2)}(\widetilde{B}) = 1 \cdot \lambda^{(1)}(B) = \lambda^{(1)}(B);$$

$$\lambda^{(2)}(\widetilde{A} \cap \widetilde{B}) = \lambda^{(1)}(A) \cdot \lambda^{(1)}(B) = \lambda^{(2)}\widetilde{A} \cap \lambda^{(2)}\widetilde{B}$$

so $\{\widetilde{A}, \widetilde{B}\}$ are independent.

Example: Let $\Omega_1 = \{1, \ldots, n\}$, $P(i) = p_i$, $p_1 + \ldots + p_n = 1$, $p_i \geq 0$, $\Omega_2 = \{1, \ldots, m\}$, $P(i) = q_i$, $q_1 + \ldots + p_n = 1$, $q_i \geq 0$ and $\Omega = \Omega_1 \times \Omega_2$, $A \subseteq \Omega : P(A) = \sum_{i,j \in A}(p_i \cdot q_j)$.

- $A \subset \Omega_1$, $B \subseteq \Omega_2$, $\widetilde{A} = A \times \Omega_2$, $\Omega_1 \times \widetilde{B}$;
- $P(\widetilde{A}) = P(A)$

$$P(\widetilde{A}) = \sum_{i,j \in \widetilde{A}} p_i q_j = \sum_{\substack{i \in A \\ j \in \Omega_2}} = \sum_{i \in A} \sum j \in \Omega_2 (p_i \cdot q_j) = \sum_{i \in A} p_i (\sum_{j \in \Omega_2} q_j) = \sum_{i \in A} p_i = P(A);$$

- $P(\widehat{B}) = P(B)$ (analogously to the previous one);
- $P(\widetilde{A} \cap \widetilde{B}) = P(\widetilde{A}) \cdot P(\widetilde{B})$

$$P(\widetilde{A} \cap \widetilde{B}) = \sum_{i,j \in \widetilde{A} \times \widetilde{B}} p_i q_j = \sum_{\substack{i \in A \\ j \in B}} p_i q_j = \sum_{i \in A} \sum_{j \in B} p_i q_i = \sum_{i \in A} p_i (\sum_{j \in B} q_j) = \sum_{i \in A} p_i P(B) =$$
$$= P(A) \cdot P(B) = P(\widetilde{A}) \cdot P(\widetilde{B});$$

so $\widetilde{A}, \widetilde{B}$ are independent.

Definition 2. *Let $A = \{A_1, \ldots, A_n\} \subseteq \mathcal{S}$. We say A is independent if for any $1 \leq i_1 < \ldots < i_k \leq n$ $(k \geq 2)$ we have:*

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k}).$$

Example: Let $\mathcal{A} = \{A, B, C\}$.

$$\mathcal{A} \text{ is independent} \equiv \begin{cases} P(A \cap B) = P(A) \cdot P(B); \\ P(A \cap C) = P(A) \cdot P(C); \\ P(B \cap C) = P(B) \cdot P(C); \\ P(A \cap B \cap C) \cdot P(A) \cdot P(B) \cdot P(C). \end{cases}$$

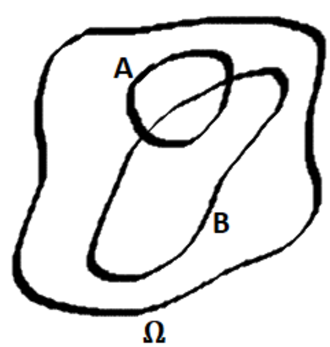
First three conditions are not enough to be certain about independence. Without the last one A, B, C are only pairwise independent.

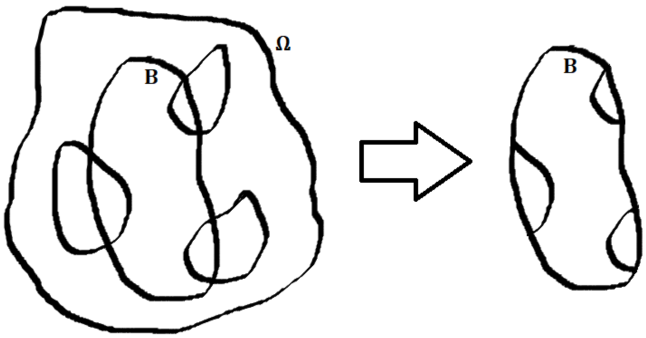
Example: Kolmogorov Cube.

3 Conditional probability

When we want to measure a probability of some particular event, knowing that another one has already occurred we count the conditional probability:

$$P(A|B) = \frac{P \cap B}{P(B)} \quad \text{where} \quad P(B) > 0.$$





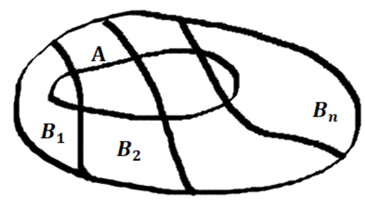
$$P(B^c|B) = \frac{P(B^c \cap B)}{P(B)} = 0$$

$$\widetilde{\mathcal{S}} = \{A \cap B : A \in \mathcal{S}\}$$

Check: $(B, \widetilde{\mathcal{S}}, \widetilde{P}) \longleftarrow$ probability space

$$\widetilde{P}(X) = \frac{P(X)}{P(B)}$$

Lets consider such Ω:



$$\Omega = \bigcup_{k=1}^{n} B_k, B_i \cap B_j = \emptyset, i \neq j \quad \text{where} \quad P(B_i) \neq 0.$$

Now let's count the probability of event A.

$$P(A) = \sum_{k=1}^{n} P(A \cap B_k) = \sum_{k=1}^{n} \frac{P(A \cap B_k)}{P(B_k)} P(B_k) = \sum_{k=1}^{n} P(A|B_k)P(B_k)$$

Fact 3. *If A and B are independent, then we have:*

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

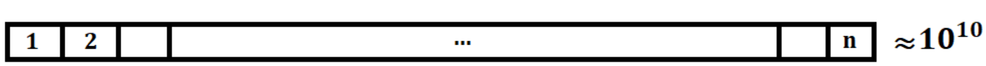
Example: Suppose that we have a dice. We throw it 2 times independently. X_i is the result of i-th throw.

$$P(1) = \ldots = P(6) = \frac{1}{6} \quad \left(\equiv \Omega = \{1, \ldots, 6\}, P(i) = \frac{1}{6}\right);$$

- $P((X_1 \leq 2) \wedge (X \geq 3)) = P(X_1 \leq 2) \cdot P(X_2 \geq 3) = \frac{2}{6} \cdot \frac{4}{6} = \frac{2}{9}$;
- $P(X_1 + X_2 = 4) = \sum_{k=1}^{3} P(X_1 + X_2 = 4 | X_1 = k) \cdot P(X_1 = k) = \frac{1}{6} P(X_1 + X_2 = 4 | X_1 = k) = \frac{1}{6} P(X_2 = 4 - k) = \frac{1}{6} \sum_{k=1}^{3} \frac{1}{6} = \frac{1}{6} \cdot \frac{3}{6} = \frac{1}{12}$.

4 Streaming

Now let's consider quite large dataset of a size about 10^{10} for which we have a pointer.



$P(X = i) = \frac{1}{n}$; where $i = 1, \ldots, n$,

```
X=0
data=NIL
n=0

onGet(A){
  n++;
  if (random()<= 1/n){
    X=n
    data=A
    }
  }
```

Let us assume that after reading n-th item we have:

$$\forall_{i\in\{1,\ldots,n\}}\left(P(X_n=i)=\frac{1}{n}\right).$$

What happens after reading $n+1$ item?

$$P\left(X_{n+1}=i\right)=P\left(X_{n+1}=i|rand()\leq\tfrac{1}{n+1}\right)\cdot P\left(rand()\leq\tfrac{1}{n+1}\right)$$
$$+\,P\left(X_{n+1}=i|rand()>\tfrac{1}{n+1}\right)\cdot\underbrace{P\left(rand()>\frac{1}{n+1}\right)}_{1-\frac{1}{n+1}}$$

$$P\left(X_{n+1}=i|rand()\leq\frac{1}{n+1}\right)\cdot\frac{1}{n+1}=\begin{cases}i\neq n+1 & \rightsquigarrow \quad 0;\\ i=n+1 & \rightsquigarrow \quad \frac{1}{n+1};\end{cases}$$

$$P\left(X_{n+1}=i|rand()>\frac{1}{n+1}\right)\cdot\left(1-\frac{1}{n+1}\right)=\begin{cases}i\neq n+1 & \rightsquigarrow \quad 0;\\ i\leq n & \rightsquigarrow \quad \frac{1}{n}\left(1-\frac{1}{n+1}\right);\end{cases}$$

$$\frac{1}{n}\left(1-\frac{1}{n+1}\right)=\frac{1}{n}\cdot\frac{n+1-1}{n+1}=\frac{n}{n(n+1)}=\frac{1}{n+1}$$

Therefore similarly to n-th item we obtain:

$$\forall_{i\in\{1,\ldots,n+1\}}\left(P\left(X_{n+1}=i\right)=\frac{1}{n+1}\right).$$