Notes on Baldi chapter 1 [1]

1 Basic Properties of Probability

Definition. A a probability measure on $(\Omega, \mathcal{F})$ is a function $P\colon \mathcal{F} \to \mathbb{R}_+$ which is normalized and σ-additive, that is $P(\Omega) = 1$ and $P\left(\cup_{n\geq 1} A_n\right) = \sum_{n\geq 1} P\left(A_n\right)$ for disjoint countable families in $\mathcal{F}$.

Recall that $A \setminus B \equiv A \cap B^c$. We will abbreviate $A \cap B$ as AB whenever not confusing.

Proposition. *Let P be a probability measure on $(\Omega, \mathcal{F})$. Then*

1. $P(\emptyset) = 0$

2. P is finitely additive: $P\left(\cup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} P\left(A_n\right)$ for any finite *disjoint family in $\mathcal{F}$, and in particular $P\left(A^c\right) = 1 - P(A)$*

3. $P(A) = P(AB) + P\left(AB^c\right)$ for all $A, B \in \mathcal{F}$

4. P is monotone: $P(A) \leq P(B)$ if $A \subseteq B$, and in particular $P(A) \leq 1$ for all $A \in \mathcal{F}$.

5. $P(A \cup B) = P(A) + P(B) - P(AB)$ for all $A, B \in \mathcal{F}$

6. P is continuous: if $\left(A_n\right)_{n\geq 1}$ is an increasing sequence (that is $A_n \subseteq A_{n+1}$ for all n) with union $\cup_{n\geq 1} A_n = A$, or it is a decreasing sequence ($A_{n+1} \subseteq A_n$ for all n) with intersection $\cap_{n\geq 1} A_n = A$, then

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

Proof. 1. Since the sequence with $A_n = \emptyset$ for all n is a disjoint family and in this case $\cup_{n\geq 1} A_n = \emptyset$, by σ-additivity we have $P(\emptyset) = \sum_{n\geq 1} P(\emptyset)$, which can be true only if $P(\emptyset) = 0$.

For 2 it suffices to take $A_n = \emptyset$ for all $n > N$. For the second assertion in 2 simply observe that $\Omega = A \cup A^c$, a disjoint union, and that $P(\Omega) = 1$.

3. $\Omega = B \cup B^c$ so $A = A\Omega = AB \cup AB^c$, disjoint union, therefore 2 implies the result.

4. If $A \subseteq B$ then $BA = A$ whence $B = A \cup BA^c$, another disjoint union; this and non-negativity of P give the result. The second assertion follows from $A \subseteq \Omega$.

5. $A \cup B = AB^c \cup AB \cup BA^c$ (disjoint), but from 3 we have $P(A) = P(AB) + P\left(AB^c\right)$ and $P(B) = P(AB) + P\left(BA^c\right)$; substituting the values of $P\left(AB^c\right)$ and $P\left(BA^c\right)$ we get the result.

6. Consider first an increasing sequence. Obviously $A_2 = A_2 A_1 \cup A_2 A_1^c = A_1 \cup (A_2 \setminus A_1)$ (disjoint) and by an easy induction argument we get $A_n = \cup_{i=1}^{n} \left(A_i \setminus A_{i-1}\right)$, with $A_0 \equiv \emptyset$, union again disjoint. We also have $A = \cup_{i\geq 1} \left(A_i \setminus A_{i-1}\right)$, disjoint; indeed obviously $\cup_{i\geq 1} \left(A_i \setminus A_{i-1}\right) \subseteq \cup_{i\geq 1} A_i$; on the other hand suppose $\omega \in A$ and let k be the smallest integer such that $\omega \in A_k$; then $\omega \in A_k \setminus A_{k-1} \subseteq \cup_{i\geq 1} \left(A_i \setminus A_{i-1}\right)$. Using σ-additivity we then

[1] S Modica 2024

conclude that

$$
\begin{aligned}
P(A) &= P\left(\cup_{i\geq 1}\left(A_i \setminus A_{i-1}\right)\right) = \sum\nolimits_{i\geq 1} P\left(A_i \setminus A_{i-1}\right) = \lim_{n\to\infty} \sum\nolimits_{i=1}^{n} P\left(A_i \setminus A_{i-1}\right) \\
&= \lim_{n\to\infty} P\left(\cup_{i=1}^{n}\left(A_i \setminus A_{i-1}\right)\right) = \lim_{n\to\infty} P(A_n).
\end{aligned}
$$

For a decreasing sequence observe that $A^c = \cup_{n\geq 1} A_n^c$ and $(A_n^c)_{n\geq 1}$ is an increasing sequence; then from what we just showed $1 - P\left(A_n\right) = P\left(A_n^c\right) \to P\left(A^c\right) = 1 - P(A)$, which implies the last result. □

The fact that $P(A) = P(AB) + P\left(AB^c\right)$ will be used all the time. The continuity of probability is central to most developments of probability theory.

Remark. Observe that when Ω is finite, $\Omega = \{\omega_1, \ldots, \omega_n\}$, to define a probability P on Ω (with $\mathcal{F} = 2^{\Omega}$) all we need to do is specify n non-negative numbers $p_i = P(\{\omega_i\})$ with $\sum_{i=1}^{n} p_i = 1$. Then for any $A \subseteq \Omega$ we have $P(A) = \sum_{i:\, \omega_i \in A} p_i$.

2 Conditional Probability and Independence

The definition of conditional probability is, with $P(B) > 0$,

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

The natural definition of independence is the following:

Definition. Let $A, B \in \mathcal{F}$. Event A is independent of B if $P(A|B) = P(A)$.

This says that knowing that B occurred does not influence the probability of A. If this is true, could it be that - asymmetrically - knowledge of A does influence the probability of B? The answer is no:

Proposition. *The following are equivalent:*

(i) A is independent of B

(ii) B is independent of A, that is $P(B|A) = P(B)$

(iii) $P(AB) = P(A) \cdot P(B)$

Proof. The equivalence $(i) \iff (iii)$ follows immediately from the definition, since

$$
\frac{P(AB)}{P(B)} = P(A) \iff P(AB) = P(A) \cdot P(B)
$$

and from this the other equivalence also follows immediately. □

Thus the concept of independence of two event is really symmetrical, and we may use as alternative definition "A and B are independent if $P(AB) = P(A)P(B)$".