Cheat Sheet: Definitions for Collections of Random Vectors

This cheat sheet organizes definitions and their associated properties from the text. Each entry is designed to be a self-contained reference for proofs.

1. Independence of Sub-σ-fields

Definition 11.1

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A finite collection of sub-σ-fields $\mathcal{A}_1, \dots, \mathcal{A}_n$ of $\mathcal{F}$ are said to be **independent** if for any choice of events $A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$, the following holds:

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot \dots \cdot P(A_n)$$

Properties and Remarks

- **Independence of Sub-collections (Remark 11.2):** If $\mathcal{A}_1, \dots, \mathcal{A}_n$ are independent, then any sub-collection is also independent. For instance, for $1 \le k_1 < \dots < k_j \le n$:

$$P(A_{k_1} \cap \dots \cap A_{k_j}) = P(A_{k_1}) \cdot \dots \cdot P(A_{k_j})$$

  - **Proof Intuition:** This follows because for any σ-field $\mathcal{A}_i$, the whole space $\Omega \in \mathcal{A}_i$. We can select Ω for the fields not in our sub-collection, and since $P(\Omega) = 1$, the equality holds.
- **Generalization to Arbitrary Collections (Definition 11.5(i)):** An arbitrary collection of sub-σ-fields $\{\mathcal{A}_i : i \in I\}$ is called **independent** if every finite sub-collection is independent.

2. Independence of Events

Definition (from Remark 11.1)

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A collection of events $A_1, \dots, A_n$ in $\mathcal{F}$ are said to be **independent** if the sub-σ-fields they generate, $\sigma(A_1), \dots, \sigma(A_n)$, are independent.

Properties and Remarks

- **Generated σ-field (Remark 11.1):** The σ-field generated by a single event A_i is $\sigma(A_i) = \{\emptyset, A_i, A_i^c, \Omega\}$.

- **Practical Condition:** The definition implies that for any choice from $\{A_i, A_i^c\}$, the joint probability is the product of individual probabilities. For example, $P(A_1 \cap A_2^c) = P(A_1)P(A_2^c)$.
- **Generalization to Arbitrary Collections (Definition 11.5(ii)):** An arbitrary collection of events $\{A_i : i \in I\}$ is **independent** if the collection of generated σ-fields $\{\sigma(A_i) : i \in I\}$ is independent.

3. Independence of Random Vectors

Definition 11.2

Let $X_1, \dots, X_n$ be n random vectors defined on a common probability space $(\Omega, \mathcal{F}, P)$, where $X_i : \Omega \to \mathbb{R}^{k_i}$. They are said to be **independent** if the sub-σ-fields they generate, $\sigma(X_1), \dots, \sigma(X_n)$, are independent.

Properties and Remarks

- **Generated σ-field:** The σ-field generated by a random vector X_i is $\sigma(X_i) = \{X_i^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^{k_i})\}$, where $\mathcal{B}(\mathbb{R}^{k_i})$ is the Borel σ-field on $\mathbb{R}^{k_i}$.
- **Equivalent Condition (Remark 11.3):** $X_1, \dots, X_n$ are independent if and only if for any choice of Borel sets $B_1 \in \mathcal{B}(\mathbb{R}^{k_1}), \dots, B_n \in \mathcal{B}(\mathbb{R}^{k_n})$:

  $$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdot \dots \cdot P(X_n \in B_n)$$

  For random variables ($k_i = 1$), this implies $P(X_1 \le t_1, \dots, X_n \le t_n) = \prod_{i=1}^n P(X_i \le t_i)$ for all $t_i \in \mathbb{R}$.
- **Product Law (Proposition 11.1):** $X_1, \dots, X_n$ are independent if and only if the law of the concatenated vector $X = (X_1, \dots, X_n)$ is the product measure of the individual laws:

  $$P_X = P_{X_1} \otimes \dots \otimes P_{X_n}$$

- **Product of PDFs (Proposition 11.4):**
  i. (**Independence $\Rightarrow$ Factorization**): If $X_1, \dots, X_n$ are independent and each X_i has a probability density function (PDF) $\phi_i(x_i)$, then the joint vector $X = (X_1, \dots, X_n)$ has a PDF $\phi(x)$ that is the product of the individual PDFs:

  $$\phi(x_1, \dots, x_n) = \prod_{i=1}^n \phi_i(x_i)$$

ii. (**Factorization $\Rightarrow$ Independence**): If the joint PDF of a vector $X = (X_1, \dots, X_n)$ can be factored as $\phi(x) = \prod_{i=1}^n \phi_i(x_i)$ where each ϕ_i is non-negative and measurable, then $X_1, \dots, X_n$ are independent.

- **Product of PMFs (Remark 11.6 for discrete RVs):** Discrete random variables $X_1, \dots, X_n$ are independent if and only if for any values $x_1, \dots, x_n$ in their respective supports:

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdot \dots \cdot P(X_n = x_n)$$

- **Expectation of Products (Proposition 11.2):** If $X_1, \dots, X_n$ are independent and $f_i : \mathbb{R}^{k_i} \to \mathbb{R}$ are measurable functions, then:

$$E\left[\prod_{i=1}^n f_i(X_i)\right] = \prod_{i=1}^n E[f_i(X_i)]$$

  This holds if either all f_i are non-negative or if $E[|f_i(X_i)|] < \infty$ for all i.
  - **Direct Consequence:** $E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$.
- **Independence vs. Uncorrelated (Proposition 11.3 & Remark 11.5):**
  - If two random variables X and Y are independent (with finite variances), they are **uncorrelated**, i.e., $\mathrm{Cov}(X, Y) = 0$.
  - The converse is **not generally true**. Uncorrelated does not imply independent. (See the $Y = GX$ example in Remark 11.5). The exception is for Gauss Vectors (see section 8).
- **Characteristic Function (Proposition 11.5):** Random variables $X_1, \dots, X_n$ are independent if and only if the joint characteristic function of the vector $X = (X_1, \dots, X_n)$ is the product of the individual characteristic functions:

$$\Phi_X(v_1, \dots, v_n) = \prod_{i=1}^n \Phi_{X_i}(v_i)$$

- **Generalization to Arbitrary Collections (Definition 11.5(iii)):** An arbitrary collection of random vectors $\{X_i : i \in I\}$ is **independent** if every finite sub-collection is independent.
- **I.I.D. (Definition 11.5(iv)):** A collection of random vectors $\{X_i : i \in I\}$ is **independent and identically distributed (i.i.d.)** if they are independent and all have the same law (i.e., $P_{X_i} = P_{X_j}$ for all $i, j \in I$).

4. n-tuple of Random Vectors

Definition 11.3

Let $X_1, \dots, X_n$ be n random vectors where $X_i : \Omega \to \mathbb{R}^{k_i}$. The **n-tuple of random vectors** is the function $X = (X_1, \dots, X_n)$ which maps from Ω to the product space $\mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_n}$. The **law of X** is the probability measure P_X defined on the product σ-field $\mathcal{B}(\mathbb{R}^{k_1}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R}^{k_n})$ by:

$$P_X(B) = P(X \in B) \quad \text{for } B \in \mathcal{B}(\mathbb{R}^{k_1}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R}^{k_n})$$

Properties and Remarks

- **Measurability (Remark 11.4):** The tuple $X = (X_1, \dots, X_n)$ is a measurable function from $(\Omega, \mathcal{F})$ to $(\mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_n}, \mathcal{B}(\mathbb{R}^{k_1}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R}^{k_n}))$. Thus, X is itself a random vector.

5. Independence of Tuples of Random Vectors

Definition 11.4

Let $Y_1, \dots, Y_N$ be a collection of N tuples of random vectors. They are said to be **independent** if the σ-fields they generate, $\sigma(Y_1), \dots, \sigma(Y_N)$, are independent.

Properties and Remarks

- **Functions of Independent Vectors (Proposition 11.6):** This is a powerful "grouping" property. If $X_1, \dots, X_n$ are independent random vectors:
    i. Any partitioning of these vectors into non-overlapping tuples results in a collection of independent tuples. For example, $Y_1 = (X_1, \dots, X_{n_1})$ and $Y_2 = (X_{n_1+1}, \dots, X_n)$ are independent.
    ii. If $T_1 = f_1(Y_1), \dots, T_p = f_p(Y_p)$ are random variables created by applying measurable functions to these independent tuples, then $T_1, \dots, T_p$ are independent.
    - **Example:** If X_1, X_2, X_3 are independent random variables, then $T_1 = f_1(X_1) = X_1^2$ and $T_2 = f_2(X_2, X_3) = X_2 + X_3$ are independent.
- **Generalization to Arbitrary Collections (Definition 11.5(v)):** An arbitrary collection of tuples $\{X_i : i \in I\}$ is **independent** if every finite sub-collection is independent.
- **I.I.D. (Definition 11.5(vi)):** This collection is **i.i.d.** if they are independent and all have the same law.

6. Law of Sums of Independent Random Vectors (Convolution)

Definition 11.6

Let $X_1, \dots, X_n$ be random vectors in $\mathbb{R}^k$. The **convolution** of their laws, denoted $P_{X_1} * \dots * P_{X_n}$, is a measure on $\mathcal{B}(\mathbb{R}^k)$ defined as the pushforward measure of the product measure $P_{X_1} \otimes \dots \otimes P_{X_n}$ by the summation map $s(x_1, \dots, x_n) = \sum_{i=1}^n x_i$:

$$(P_{X_1} * \dots * P_{X_n})(B) := (P_{X_1} \otimes \dots \otimes P_{X_n})(s^{-1}(B))$$

Properties and Remarks

- **Law of the Sum (Proposition 11.7(i)):** If $X_1, \dots, X_n$ are **independent** random vectors in $\mathbb{R}^k$, then the law of their sum $Z = \sum_{i=1}^n X_i$ is the convolution of their individual laws:

$$P_Z = P_{X_1} * \dots * P_{X_n}$$

- **Convolution of PDFs (Proposition 11.7(ii)):** If additionally each X_i has a PDF ϕ_i, then the sum Z has a PDF ϕ_Z given by the n-fold convolution of the individual PDFs:

$$\phi_Z(z) = (\phi_1 * \dots * \phi_n)(z)$$

- **Sum of Discrete RVs (Remark 11.10):** If $X_1, \dots, X_n$ are independent discrete random vectors with supports $E_1, \dots, E_n$ and PMFs $p_1, \dots, p_n$, the PMF of the sum $Z = \sum X_i$ is given by:

$$P_Z(\{z\}) = \sum_{x_n \in E_n} \dots \sum_{x_2 \in E_2} p_1(z - x_2 - \dots - x_n) p_2(x_2) \cdots p_n(x_n)$$

  For two variables, this simplifies to: $P_{X_1+X_2}(\{z\}) = \sum_{x_2 \in E_2} P_{X_1}(\{z - x_2\}) P_{X_2}(\{x_2\})$.
- **Characteristic Function of a Sum (Proposition 11.8):** If $X_1, \dots, X_n$ are independent, the characteristic function of the sum $Z = \sum X_i$ is the product of the individual characteristic functions:

$$\Phi_Z(v) = \prod_{i=1}^n \Phi_{X_i}(v)$$

  - **Usefulness:** This is often much easier to work with than performing convolutions. For example, it easily shows that the sum of independent Normal random variables is Normal.

7. Gauss Vector (Multivariate Normal)

Definition 11.7

A random vector $X = (X_1, \ldots, X_k)$ is a **Gauss vector** if for any vector of constants $v = (v_1, \ldots, v_k) \in \mathbb{R}^k$, the linear combination $v^t X = \sum_{i=1}^k v_i X_i$ is a (univariate) Gaussian random variable.

Properties and Remarks

- **Notation (Definition 11.8):** We write $X \sim \mathcal{N}(\mu, \Sigma)$ to denote a Gauss vector with mean vector $\mu = E[X]$ and covariance matrix Σ.
- **Components are Gaussian (Remark 11.11(i)):** If X is a Gauss vector, then each component X_i is a Gaussian random variable. (To see this, choose $v = e_i$, the i-th standard basis vector).
- **Converse is False (Remark 11.11(ii)):** A vector whose components are all Gaussian is **not necessarily** a Gauss vector. The sum of the components may not be Gaussian.
- **Independent Gaussians form a Gauss Vector (Remark 11.11(iii)):** If $X_1, \ldots, X_k$ are *independent* Gaussian random variables, then the vector $X = (X_1, \ldots, X_k)$ *is* a Gauss vector.
- **Characteristic Function (Proposition 11.9):** The characteristic function of a Gauss vector $X \sim \mathcal{N}(\mu, \Sigma)$ is given by:

  $$\Phi_X(v) = \exp\left(i\mu^t v - \frac{1}{2} v^t \Sigma v\right)$$

- **Independence Condition (Proposition 11.10) - CRITICAL PROPERTY:** Let $X \sim \mathcal{N}(\mu, \Sigma)$ be a Gauss vector. The components $X_1, \ldots, X_k$ are **independent** if and only if they are pairwise uncorrelated, which is equivalent to the covariance matrix Σ being a **diagonal matrix**.

  $$X_1, \ldots, X_k \text{ independent} \iff \text{Cov}(X_i, X_j) = 0 \text{ for all } i \neq j \iff \Sigma \text{ is diagonal.}$$

  This provides a simple way to check for independence that only works for Gauss vectors.
- **PDF (Proposition 11.11):** If the covariance matrix Σ is positive definite (and thus invertible), the Gauss vector X has a a multivariate normal probability density function given by:

  $$\phi(x) = \frac{1}{\sqrt{(2\pi)^k \det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)\right)$$

- **Standard Gauss Vector (Definition 11.8):** A Gauss vector is called **standard** if its mean is the zero vector ($\mu = 0$) and its covariance matrix is the identity matrix ($\Sigma = I$).

8. Conditional Probability and Expectation

Definition (Sec 11.4)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $B \in \mathcal{F}$ be an event with $P(B) > 0$.

1. The **conditional probability given B** is the measure P_B on $(\Omega, \mathcal{F})$ defined by:

$$P_B(A) := P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{for any } A \in \mathcal{F}$$

2. The **conditional expectation of a random variable X given B** is defined as:

$$E[X|B] = \frac{E[X \cdot 1_B]}{P(B)}$$

Properties and Remarks

- P_B is a valid probability measure on $\mathcal{F}$, satisfying the axioms of probability.
- The conditional expectation $E[X|B]$ is simply the standard expectation of X with respect to the new conditional probability measure P_B.

$$E[X|B] = \int_\Omega X(\omega)\, dP_B(\omega)$$