Lecture 3 / Week 2

Random Vectors, Distributions and Integrals

Definition Suppose we have probability space $(\Omega, \mathcal{F}, P)$: A **random vector** is a function $X : \Omega \to \mathbb{R}^K$ that is $\mathcal{F}\backslash\mathcal{B}(\mathbb{R}^K)$ measurable.

X is a K-dimensional random variable: $X(\omega) = (X_1(\omega), X_2(\omega)...., X_K(\omega))$, so $X = (X_1, X_2...., X_K)$ is measurable. Also note that X is $\mathcal{F}\backslash\mathcal{B}(\mathbb{R}^K)$ measurable if and only if $X_1, X_2...., X_K$ are $\mathcal{F}\backslash\mathcal{B}(\mathbb{R})$ measurable. So, it is sufficient that the coordinates of the vector are measurable. This amounts to saying that $X = (X_1, X_2...., X_K)$ is a random variable iff $X_1, X_2...., X_K$ are random variables.

Definition If we have a class $\mathcal{G} \subset \mathcal{F}$ (a sub-σ-algebra of $\mathcal{F}$) then X is a $\mathcal{G}\backslash\mathcal{B}(\mathbb{R}^K)$ **measurable** if and only if $X_1, X_2...., X_K$ are $\mathcal{G}\backslash\mathcal{B}(\mathbb{R})$ measurable. For notational simplicity we will say X is $\mathcal{G}$ measurable if and only if $X_1, X_2...., X_K$ are $\mathcal{G}$ measurable.

Definition The **smallest σ-algebra** w.r.t which a random vector is measurable is denoted by $\sigma(X) = \cap_{X \text{ is } \mathcal{G}\ measurable}\mathcal{G}$ is a σ-algebra. By definition $X^{-1}(B) \in \mathcal{G}$. [X is $\mathcal{G}$ measurable if $X^{-1}(B) \in \mathcal{G}$]. But then, $X^{-1}(B) \in \sigma(X), i.e.$ X is $\sigma(X)$ measurable. Note that if we take the class of inverse images of Borel sets, this class will be contained in $\sigma(X)$, in fact it will be equal to $\sigma(X)$. Formally $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^K)\}$. We already said that $\sigma(X) \supset \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^K)\}$, since $X^{-1}(B) \in \sigma(X)$, to show the reverse inclusion first we need to show that the class $\{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^K)\}$ is a σ-algebra. If we take a class $\mathcal{G} \subset \mathcal{F}$ such that X is $\mathcal{G}$ *measurable* then $\sigma(X) \subset \mathcal{G}$. If $\sigma(X) \subset \mathcal{G} \Rightarrow X^{-1}(B) \in \mathcal{G} \Rightarrow X$ is $\mathcal{G}$ *measurable*. Hence X is $\mathcal{G}$ *measurable* iff $\sigma(X) \subset \mathcal{G}$.

Exercise Show that the class $\{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^K)\}$ is a σ-algebra.

Proof To show that the class $\{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^K)\}$ (Lets call it C) is a σ-algebra, we have to check the 3 properties of σ-algebra. 1.) By definition of random variable function $X^{-1}(\mathbb{R}^K) = \Omega, \mathbb{R}^K \in \mathcal{B}(\mathbb{R}^K)$, so $\Omega \in C$.2.) If $X^{-1}(B) \in C$, then $(X^{-1}(B))^C \in C$ should hold. But $X^{-1}(B))^C = X^{-1}(B^C)$, since $\mathcal{B}(\mathbb{R}^K)$ is a σ-algebra,$B^C \in \mathcal{B}(\mathbb{R}^K)$, $X^{-1}(B^C) \in C$. 3.)If $X^{-1}(B_1), X^{-1}(B_2).... \in C$, then $\cup_{i=1}^{\infty} X^{-1}(B_i) \in C$. This condition holds, since $\mathcal{B}(\mathbb{R}^K)$ is σ-algebra, then $\cup_{i=1}^{\infty} B_i \in \mathcal{B}(\mathbb{R}^K)$, but then $\cup_{i=1}^{\infty} X^{-1}(B_i) \in C$. QED.

Definition Suppose we have probability space $(\Omega, \mathcal{F}, P)$ and a random vector $X : (\Omega \to \mathbb{R}^K, on\ (\mathbb{R}^K, \mathcal{B}(\mathbb{R}^K))$. Then μ_X is called the **probability distribution** of X. We take a Borel set $B \in \mathcal{B}(\mathbb{R}^K)$ and define its measure $\mu_X(B) = P(X^{-1}(B)) = P(X \in B)$. It is the measure of the probability of all ω s.t $\{\omega \in \Omega : X(\omega) \in B\}$

Insert here Figure 1

Claim μ_X is a probability measure on $(\mathbb{R}^K, \mathcal{B}(\mathbb{R}^K))$.Then we have the new probability space $(\Omega, \mathcal{F}, \mu_X)$.

Proof To prove that μ_X is a probability measure, we should check the 3 properties of probability measure. 1.)$\mu_X(B) \geq 0$, by definition and the fact that P is a probability $\mu_X(B) = P(X^{-1}(B)) \geq 0$. 2.) $\mu_X(\mathbb{R}^K) = 1$: By definition $\mu_X(\mathbb{R}^K) = P(X^{-1}(\mathbb{R}^K)) = P(\Omega) = 1$. 3.) $\mu_X(\cup_{j=1}^{\infty} B_j) = \sum_{j=1}^{\infty} \mu_K(B_j), B_i \cap B_j = \emptyset$, then by definition $\mu_X(\cup_{j=1}^{\infty} B_j) = P(X^{-1}(\cup_{j=1}^{\infty} B_j)) = P(\cup_{j=1}^{\infty} X^{-1}(B_j))$, since P is a probability $P(\cup_{j=1}^{\infty} X^{-1}(B_j)) = \sum_{j=1}^{\infty} P(X^{-1}(B_j)) = \sum_{j=1}^{\infty} \mu_K(B_j)$.QED

Insert here Figure 2

Distribution Function

$$\mathrm{F}_X : \mathbb{R}^K \to \mathbb{R}$$
$$\mathrm{F}_X(x_1, x_2, ...x_K) = P(X_1 \leq x_1, X_2 \leq x_2, ..., X_K \leq x_K)$$
$$\mathrm{F}_X(x_1, x_2, ...x_K) = \mu_X((-\infty, x_1]x(-\infty, x_2]x...(-\infty, x_K])$$

Example If X is a random variable, $\mathrm{F}_X(x) = P(X \leq x) = P(X \in (-\infty, x]) = \mu_X((-\infty, x])$. If $X = (X_1, X_2)$, $\mathrm{F}_X(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = P(X \in (-\infty, x_1]x(-\infty, x_2]) = \mu_X((-\infty, x_1]x(-\infty, x_2])$.

Insert here Figure 3

Homework Review properties of distribution functions!

Probability Distributions

Suppose we have probability space $(\Omega, \mathcal{F}, P)$ and a random vector $X = (X_1, X_2...., X_K)$, then μ_X on $\mathcal{B}(\mathbb{R}^K)$ is a probability distribution of X.

Insert here Figure 4

Definition μ_X (B)=$\sum_{s \in B} m(s)$; μ_X is **discrete** if there exists a countable (at most) set $S = \{s_1, s_2.....\}$ in $\mathbb{R}^K$ such that

$$\begin{aligned} \mu_X(B) &= \sum_{i=1}^{\infty} m(s_i) \mathbf{1}_B(s_i) = \sum_{i=1}^{\infty} m(s_i) \delta_{s_i}(B) \\ where\ \delta_{s_i}(B) &= \left\{ \begin{array}{cc} 1 & s_i \in B \\ 0 & s_i \notin B \end{array} \right\} is\ \textbf{Dirac Measure}. \end{aligned}$$

$\delta_s(A)$ is the probability measure that puts all its mass on s. Note that $\mathbf{1}_B(s_i) = \delta_{s_i}(B)$. The above equation can also be read as μ_X is a convex combination of Dirac probability measures. (convex$\Rightarrow m(s_i) \leq 1, \Sigma m(s_i) = 1$).

Definition Suppose we have probability space $(\Omega, \mathcal{F}, P)$ and a random vector $X = (X_1, X_2 ...., X_K)$ with probability distribution μ_X on $\mathcal{B}(\mathbb{R}^K)$. The probability distribution μ_X is called **absolutely continuous** if there exists a nonnegative integrable function f, (**density function** of X), such that the **distribution function** F_X (equivalently μ_X) can be written as

$$\boxed{\mu_X(B){=}\int_B f(x)dx}$$

$$\begin{aligned} & \textit{where } B \text{ is a rectangle.} \\ B &= (a_1, b_1]x(a_2, b_2]x....(a_K, b_K] \\ \int_B f(x)dx &= \int_{a_1}^{b_1}\int_{a_2}^{b_2} .... \int_{a_K}^{b_K} f(x_1, ...x_K)\ dx_1...d_K \end{aligned}$$

Riemann Integral

I= rectangle, $f : \mathbb{R}^K \to \mathbb{R}$, f is continuous

Definition A collection of disjoint rectangles such that their union is I is called **partition** of I

$$\begin{aligned} \int_I f(x)dx &= \sup \sum_{j=1}^{n} (\inf_{x \in I_j} f(x)).\lambda(I_j) \\ & \textit{where } I \text{ is measured with} \\ \lambda(I) &= \prod_{i=1}^{K} (b_i - a_i) \end{aligned}$$

is called the **Riemann Integral**.

Insert here Figure 5

As we can see by the definition of Riemann Integral, that it is based on rectangles, so it is not suitable to integrate density functions of other shapes. So, we have to extend the the measure $\lambda(I)$ to Borel sets.

$$\lambda(B) = \inf \ \sum_{j=1}^{n} \lambda(I)_{j=1,2...n}$$

Note that the infimum is defined over all families of rectangles $(I)_{j=1,2...n}$ such that B $\subset \cup_{j=1}^{n} I_j$.

Definition A family of rectangles $(I)_{j=1,2...n}$ such that B $\subset \cup_{j=1}^{n} I_j$ is called **cover** of B.

Definition The above defined measure $\lambda(B)$ is called **Lebesgue measure**.

Insert here Figure 6

Lebesgue Integral

Now we can define the **Lebesgue Integral** on Borel sets.

$$\int_B f(x)dx = \sup \sum_{j=1}^{n} (\inf_{x \in B_j} f(x)).\lambda(B_j)$$

Note that for **Lebesgue Integral** we do not need a continuity assumption on density function, it is sufficient that the density function is measurable. Lebesgue Integral coincides with Riemann Integral if it is defined on rectangles; $\int_B f(x)dx = \int_{I=B} f(x)dx$