Stat 709: Mathematical Statistics
Lecture 2

Jun Shao

Department of Statistics
University of Wisconsin
Madison, WI 53706, USA

Lecture 2: Product measure, measurable function and distribution

Product space

$\mathscr{I} = \{1,...,k\}$, k is finite or ∞

Γ_i, $i \in \mathscr{I}$, are some sets

$\prod_{i\in\mathscr{I}} \Gamma_i = \Gamma_1 \times \cdots \times \Gamma_k = \{(a_1,...,a_k) : a_i \in \Gamma_i, i \in \mathscr{I}\}$

$\mathscr{R} \times \mathscr{R} = \mathscr{R}^2$, $\mathscr{R} \times \mathscr{R} \times \mathscr{R} = \mathscr{R}^3$

Product σ-field

Let $(\Omega_i, \mathscr{F}_i)$, $i \in \mathscr{I}$, be measurable spaces

$\prod_{i\in\mathscr{I}} \mathscr{F}_i$ is not necessarily a σ-field

$\sigma(\prod_{i\in\mathscr{I}} \mathscr{F}_i)$ is called the *product σ-field* on the *product space* $\prod_{i\in\mathscr{I}} \Omega_i$

$(\prod_{i\in\mathscr{I}} \Omega_i, \sigma(\prod_{i\in\mathscr{I}} \mathscr{F}_i))$ is denoted by $\prod_{i\in\mathscr{I}} (\Omega_i, \mathscr{F}_i)$

Example: $\prod_{i=1,...,k} (\mathscr{R}, \mathscr{B}) = (\mathscr{R}^k, \mathscr{B}^k)$

Lecture 2: Product measure, measurable function and distribution

Product space

$\mathscr{I} = \{1, ..., k\}$, k is finite or ∞
Γ_i, $i \in \mathscr{I}$, are some sets
$\prod_{i\in\mathscr{I}} \Gamma_i = \Gamma_1 \times \cdots \times \Gamma_k = \{(a_1, ..., a_k) : a_i \in \Gamma_i, i \in \mathscr{I}\}$
$\mathscr{R} \times \mathscr{R} = \mathscr{R}^2$, $\mathscr{R} \times \mathscr{R} \times \mathscr{R} = \mathscr{R}^3$

Product σ-field

Let $(\Omega_i, \mathscr{F}_i)$, $i \in \mathscr{I}$, be measurable spaces
$\prod_{i\in\mathscr{I}} \mathscr{F}_i$ is not necessarily a σ-field
$\sigma(\prod_{i\in\mathscr{I}} \mathscr{F}_i)$ is called the *product σ-field* on the *product space* $\prod_{i\in\mathscr{I}} \Omega_i$
$(\prod_{i\in\mathscr{I}} \Omega_i, \sigma(\prod_{i\in\mathscr{I}} \mathscr{F}_i))$ is denoted by $\prod_{i\in\mathscr{I}} (\Omega_i, \mathscr{F}_i)$
Example: $\prod_{i=1,...,k} (\mathscr{R}, \mathscr{B}) = (\mathscr{R}^k, \mathscr{B}^k)$

Product measure

Consider a rectangle $[a_1,b_1]\times[a_2,b_2]\subset\mathscr{R}^2$.
The usual area of $[a_1,b_1]\times[a_2,b_2]$ is

$$(b_1-a_1)(b_2-a_2)=m([a_1,b_1])m([a_2,b_2])$$

Is $m([a_1,b_1])m([a_2,b_2])$ the same as the value of a measure defined on the product σ-field?

σ-finite

A measure ν on $(\Omega,\mathscr{F})$ is said to be σ-*finite* iff there exists a sequence $\{A_1,A_2,...\}$ such that $\cup A_i=\Omega$ and $\nu(A_i)<\infty$ for all i
Any finite measure (such as a probability measure) is clearly σ-finite
The Lebesgue measure on $\mathscr{R}$ is σ-finite, since $\mathscr{R}=\cup A_n$ with $A_n=(-n,n)$, $n=1,2,...$
The counting measure in is σ-finite if and only if Ω is countable
The measure $\nu(A)=\infty$ unless $A=\emptyset$ is not σ-finite

Product measure

Consider a rectangle $[a_1,b_1]\times[a_2,b_2]\subset\mathscr{R}^2$.
The usual area of $[a_1,b_1]\times[a_2,b_2]$ is

$$(b_1-a_1)(b_2-a_2)=m([a_1,b_1])m([a_2,b_2])$$

Is $m([a_1,b_1])m([a_2,b_2])$ the same as the value of a measure defined on the product σ-field?

σ-finite

A measure ν on $(\Omega,\mathscr{F})$ is said to be σ-*finite* iff there exists a sequence $\{A_1,A_2,...\}$ such that $\cup A_i=\Omega$ and $\nu(A_i)<\infty$ for all i
Any finite measure (such as a probability measure) is clearly σ-finite
The Lebesgue measure on $\mathscr{R}$ is σ-finite, since $\mathscr{R}=\cup A_n$ with $A_n=(-n,n)$, $n=1,2,...$
The counting measure in is σ-finite if and only if Ω is countable
The measure $\nu(A)=\infty$ unless $A=\emptyset$ is not σ-finite

Proposition 1.3 (Product measure theorem)

Let $(\Omega_i, \mathscr{F}_i, \nu_i)$, $i = 1, ..., k$, be measure spaces with σ-finite measures, where $k \geq 2$ is an integer.
Then there exists a unique σ-finite measure on the product σ-field $\sigma(\mathscr{F}_1 \times \cdots \times \mathscr{F}_k)$, called the *product measure* and denoted by $\nu_1 \times \cdots \times \nu_k$, such that

$$\nu_1 \times \cdots \times \nu_k(A_1 \times \cdots \times A_k) = \nu_1(A_1) \cdots \nu_k(A_k)$$

for all $A_i \in \mathscr{F}_i$, $i = 1, ..., k$.

The *joint* c.d.f.

Let P be a probability measure on $(\mathscr{R}^k, \mathscr{B}^k)$.
The *joint* c.d.f. of P is defined by

$$F(x_1, ..., x_k) = P\left((-\infty, x_1] \times \cdots \times (-\infty, x_k]\right), \quad x_i \in \mathscr{R}$$

There is a one-to-one correspondence between probability measures and joint c.d.f.'s on $\mathscr{R}^k$

Proposition 1.3 (Product measure theorem)

Let $(\Omega_i, \mathscr{F}_i, \nu_i)$, $i = 1, ..., k$, be measure spaces with σ-finite measures, where $k \geq 2$ is an integer.
Then there exists a unique σ-finite measure on the product σ-field $\sigma(\mathscr{F}_1 \times \cdots \times \mathscr{F}_k)$, called the *product measure* and denoted by $\nu_1 \times \cdots \times \nu_k$, such that

$$\nu_1 \times \cdots \times \nu_k(A_1 \times \cdots \times A_k) = \nu_1(A_1) \cdots \nu_k(A_k)$$

for all $A_i \in \mathscr{F}_i$, $i = 1, ..., k$.

The *joint* c.d.f.

Let P be a probability measure on $(\mathscr{R}^k, \mathscr{B}^k)$.
The *joint* c.d.f. of P is defined by

$$F(x_1, ..., x_k) = P\left((-\infty, x_1] \times \cdots \times (-\infty, x_k]\right), \quad x_i \in \mathscr{R}$$

There is a one-to-one correspondence between probability measures and joint c.d.f.'s on $\mathscr{R}^k$

Marginal c.d.f.

If $F(x_1,...,x_k)$ is a joint c.d.f., then

$$F_i(x) = \lim_{x_j\to\infty, j=1,...,i-1,i+1,...,k} F(x_1,...,x_{i-1},x,x_{i+1},...,x_k)$$

is a c.d.f. and is called the ith *marginal* c.d.f.

The c.d.f. and product measure

- Marginal c.d.f.'s are determined by their joint c.d.f.
- But a joint c.d.f. cannot be determined by k marginal c.d.f.'s.
- If

  $$F(x_1,...,x_k) = F_1(x_1)\cdots F_k(x_k), \quad (x_1,...,x_k) \in \mathscr{R}^k,$$

  then the probability measure corresponding to F is the product measure $P_1 \times \cdots \times P_k$ with P_i being the probability measure corresponding to F_i

Marginal c.d.f.

If $F(x_1,...,x_k)$ is a joint c.d.f., then

$$F_i(x) = \lim_{x_j\to\infty, j=1,...,i-1,i+1,...,k} F(x_1,...,x_{i-1},x,x_{i+1},...,x_k)$$

is a c.d.f. and is called the ith *marginal* c.d.f.

The c.d.f. and product measure

- Marginal c.d.f.'s are determined by their joint c.d.f.
- But a joint c.d.f. cannot be determined by k marginal c.d.f.'s.
- If

  $$F(x_1,...,x_k) = F_1(x_1)\cdots F_k(x_k), \quad (x_1,...,x_k)\in\mathscr{R}^k,$$

  then the probability measure corresponding to F is the product measure $P_1\times\cdots\times P_k$ with P_i being the probability measure corresponding to F_i

Measurable function

f: a function from Ω to Λ (often $\Lambda = \mathscr{R}^k$)
Inverse image of $B \subset \Lambda$ under f:

$$f^{-1}(B) = \{f \in B\} = \{\omega \in \Omega : f(\omega) \in B\}.$$

The inverse function f^{-1} need not exist for $f^{-1}(B)$ to be defined.

$$f^{-1}(B^c) = (f^{-1}(B))^c \quad \text{for any } B \subset \Lambda$$

$$f^{-1}(\cup B_i) = \cup f^{-1}(B_i) \quad \text{for any } B_i \subset \Lambda,\ i = 1, 2, \ldots$$

Let $\mathscr{C}$ be a collection of subsets of Λ.
Define $f^{-1}(\mathscr{C}) = \{f^{-1}(C) : C \in \mathscr{C}\}$

Definition 1.3

Let $(\Omega, \mathscr{F})$ and $(\Lambda, \mathscr{G})$ be measurable spaces.
Let f be a function from Ω to Λ.
f is called a *measurable function* from $(\Omega, \mathscr{F})$ to $(\Lambda, \mathscr{G})$ iff $f^{-1}(\mathscr{G}) \subset \mathscr{F}$.

Measurable function

f: a function from Ω to Λ (often $\Lambda = \mathscr{R}^k$)
Inverse image of $B \subset \Lambda$ under f:

$$f^{-1}(B) = \{f \in B\} = \{\omega \in \Omega : f(\omega) \in B\}.$$

The inverse function f^{-1} need not exist for $f^{-1}(B)$ to be defined.

$$f^{-1}(B^c) = (f^{-1}(B))^c \quad \text{for any } B \subset \Lambda$$

$$f^{-1}(\cup B_i) = \cup f^{-1}(B_i) \quad \text{for any } B_i \subset \Lambda,\ i = 1, 2, \ldots$$

Let $\mathscr{C}$ be a collection of subsets of Λ.
Define $f^{-1}(\mathscr{C}) = \{f^{-1}(C) : C \in \mathscr{C}\}$

Definition 1.3

Let $(\Omega, \mathscr{F})$ and $(\Lambda, \mathscr{G})$ be measurable spaces.
Let f be a function from Ω to Λ.
f is called a *measurable function* from $(\Omega, \mathscr{F})$ to $(\Lambda, \mathscr{G})$ iff $f^{-1}(\mathscr{G}) \subset \mathscr{F}$.

Remarks

- f is measurable from $(\Omega, \mathscr{F})$ to $(\Lambda, \mathscr{G})$ iff for any $B \in \mathscr{G}$, $f^{-1}(B) = \{\omega : f(\omega) \in B\} \in \mathscr{F}$; we don't care about whether $\{f(\omega) : \omega \in A\}$ is in $\mathscr{G}$ or not, $A \in \mathscr{F}$.
- If $\mathscr{F}$ is the collection of all subsets of Ω, then any function f is measurable.
- If f is measurable from $(\Omega, \mathscr{F})$ to $(\Lambda, \mathscr{G})$, then $f^{-1}(\mathscr{G})$ is a sub-σ-field of $\mathscr{F}$ and is called the σ-field generated by f and denoted by $\sigma(f)$.
- $\sigma(f)$ may be much simpler than $\mathscr{F}$
- A measurable f from $(\Omega, \mathscr{F})$ to $(\mathscr{R}, \mathscr{B})$ is called a Borel function.
- In probability and statistics, a Borel function is also called a random variable.
  A random variable = a variable that is random?
- A random vector $(X_1, ..., X_n)$ is a function measurable from $(\Omega, \mathscr{F})$ to $(\mathscr{R}^n, \mathscr{B}^n)$.
- $(X_1, ..., X_n)$ is a random vector iff each X_i is a random variable.

Indicator functions

The indicator function for $A \subset \Omega$ is:

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A. \end{cases}$$

For any $B \subset \mathscr{R}$,

$$I_A^{-1}(B) = \begin{cases} \emptyset & 0 \notin B, 1 \notin B \\ A & 0 \notin B, 1 \in B \\ A^c & 0 \in B, 1 \notin B \\ \Omega & 0 \in B, 1 \in B. \end{cases}$$

Then, $\sigma(I_A) = \{\emptyset, A, A^c, \Omega\}$ and I_A is Borel iff $A \in \mathscr{F}$
Note that $\sigma(I_A)$ is much simpler than $\mathscr{F}$.

Simple functions

Let $A_1,...,A_k$ be measurable sets on Ω and $a_1,...,a_k$ be real numbers. A simple function is

$$\varphi(\omega) = \sum_{i=1}^{k} a_i I_{A_i}(\omega),$$

Let $A_1,...,A_k$ be a partition of Ω, i.e., A_i's are disjoint and $A_1 \cup \cdots \cup A_k = \Omega$.

Then the simple function φ with distinct a_i's exactly characterizes this partition and $\sigma(\varphi) = \sigma(\{A_1,...,A_k\})$.

A simple function is nonnegative iff $a_i \geq 0$ for all i.

Proposition 1.4

Let $(\Omega, \mathscr{F})$ be a measurable space.

(i) f is Borel if and only if $f^{-1}(a,\infty) \in \mathscr{F}$ for all $a \in \mathscr{R}$.

(ii) If f and g are Borel, then so are fg and $af + bg$, where a and b are real numbers; also, f/g is Borel provided $g(\omega) \neq 0$ for any $\omega \in \Omega$.

Simple functions

Let $A_1,...,A_k$ be measurable sets on Ω and $a_1,...,a_k$ be real numbers. A simple function is

$$\varphi(\omega) = \sum_{i=1}^{k} a_i I_{A_i}(\omega),$$

Let $A_1,...,A_k$ be a partition of Ω, i.e., A_i's are disjoint and $A_1 \cup \cdots \cup A_k = \Omega$.
Then the simple function φ with distinct a_i's exactly characterizes this partition and $\sigma(\varphi) = \sigma(\{A_1,...,A_k\})$.
A simple function is nonnegative iff $a_i \geq 0$ for all i.

Proposition 1.4

Let $(\Omega, \mathscr{F})$ be a measurable space.

(i) f is Borel if and only if $f^{-1}(a,\infty) \in \mathscr{F}$ for all $a \in \mathscr{R}$.

(ii) If f and g are Borel, then so are fg and $af + bg$, where a and b are real numbers; also, f/g is Borel provided $g(\omega) \neq 0$ for any $\omega \in \Omega$.

Proposition 1.4 (continued)

(iii) If $f_1, f_2, \ldots$ are Borel, then so are $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$, and $\liminf_n f_n$.

Furthermore, the set

$$A = \left\{ \omega \in \Omega : \lim_{n\to\infty} f_n(\omega) \text{ exists} \right\}$$

is an event and the function

$$h(\omega) = \begin{cases} \lim_{n\to\infty} f_n(\omega) & \omega \in A \\ f_1(\omega) & \omega \notin A \end{cases}$$

is Borel.

(iv) Suppose that f is measurable from $(\Omega, \mathscr{F})$ to $(\Lambda, \mathscr{G})$ and g is measurable from $(\Lambda, \mathscr{G})$ to $(\Delta, \mathscr{H})$.

Then the composite function $g \circ f$ is measurable from $(\Omega, \mathscr{F})$ to $(\Delta, \mathscr{H})$.

(v) Let Ω be a Borel set in $\mathscr{R}^p$.

If f is a continuous function from Ω to $\mathscr{R}^q$, then f is measurable.

Distribution (law)

Let $(\Omega, \mathscr{F}, \nu)$ be a measure space and f be a measurable function from $(\Omega, \mathscr{F})$ to $(\Lambda, \mathscr{G})$.
The *induced measure* by f, denoted by $\nu \circ f^{-1}$, is a measure on $\mathscr{G}$ defined as

$$\nu \circ f^{-1}(B) = \nu(f \in B) = \nu\left(f^{-1}(B)\right), \quad B \in \mathscr{G}$$

If $\nu = P$ is a probability measure and X is a random variable or a random vector, then $P \circ X^{-1}$ is called the *law* or the *distribution* of X and is denoted by P_X.
The c.d.f. of P_X is also called the c.d.f. or joint c.d.f. of X and is denoted by F_X.
Examples 1.3 and 1.4