# Math 6261 23-02-03.

## §1.5 Expected values.

Let $(\Omega, \mathcal{F}, P)$ be a prob. space. Recall that a random variable $X$ on $(\Omega, \mathcal{F}, P)$ is a measurable function $X: \Omega \to \mathbb{R}$, i.e.

$$X^{-1}(A) \in \mathcal{F} \quad \text{for each Borel set } A \subset \mathbb{R}.$$

Let $EX = \int X \, dP$ and we call it the expectation of $X$.

Notice that $EX$ is well-defined if

$$\int X^+ \, dp < \infty \quad \text{or} \quad \int X^- \, dp < \infty,$$

where $X^+ = \max\{0, X\}$, $X^- = \max\{0, -X\}$.

According to the properties of integration, we have.

Prop 1.6. Let $X, Y \geq 0$, or $E|X|, E|Y| < \infty$, then

- $E(X+Y) = E(X) + E(Y)$
- $E(aX+b) = aE(X) + b$ for all $a, b \in \mathbb{R}$.
- $EX \geq EY$ if $X \geq Y$.

Prop 1.7 (i) (Jensen Inequality) Suppose $\varphi : \mathbb{R} \to \mathbb{R}$ is convex. Then

$$E(\varphi(X)) \geq \varphi(EX)$$

if $E|X|$ and $E(|\varphi(X)|) < \infty$.

(ii) Hölder inequality

$$E|XY| \leq \|X\|_p \|Y\|_q,$$

where $p, q > 1$, $1/p + 1/q = 1$.

Here $\|X\|_p = \int |X|^p \, dP$

(iii) (Chebyshev's inequality)

Suppose $\varphi : \mathbb{R} \to \mathbb{R}_+$. Then for each Borel $A \subset \mathbb{R}$,

$$\inf\{\varphi(y) : y \in A\} \cdot P\{X \in A\} \leq \int_{\{X \in A\}} \varphi(X) \, dP$$

In particular for each $a > 0$,

$$a^2 P\{|X| \geq a\} \leq EX^2.$$

Prop 1.8 (Fatou's lemma) If $X_n \geq 0$, then

$$\liminf_{n\to\infty} EX_n \geq E\left(\liminf_{n\to\infty} X_n\right).$$

(Monotone convergence Thm) If $0 \leq X_n \uparrow X$, then

$$\lim_{n\to\infty} EX_n = EX$$

(Dominated Convergence Thm) If $X_n \to X$ a.e., $|X_n| \leq Y$, $EY < \infty$, then

$$\lim_{n\to\infty} EX_n = EX.$$

Thm 1.9 (Change of variables formula).

Let $X$ be a random element in $(T, \mathcal{T})$ with distribution $\mu$, that is,

$$\mu(A) = P\{X \in A\}, \quad A \in \mathcal{T}.$$

Let $f: (T, \mathcal{T}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be measurable such that $f \geq 0$ or $E(|f(X)|) < \infty$. Then

$$E f(X) = \int_T f(y)\, d\mu(y).$$

Pf. Case 1. $f = \mathbb{1}_A$, where $A \in \mathcal{T}$.

Then $$E f(X) = \int \mathbb{1}_A(X(\omega))\, dP(\omega)$$
$$= \int \mathbb{1}_{X^{-1}(A)}(\omega)\, dP(\omega)$$
$$= P(X^{-1}(A)) = P(X \in A)$$

$$\int \mathbb{1}_A(y)\, d\mu(y) = \mu(A) = P(X \in A).$$
$$= E f(X).$$

Case 2. $f = \sum_{i=1}^{k} \alpha_i \mathbb{1}_{A_i}$

By the linearity of integration,

$$E f(X) = \sum_{i=1}^{k} \alpha_i E(\mathbb{1}_{A_i}(X))$$
$$= \sum_{i=1}^{k} \alpha_i \int \mathbb{1}_{A_i}\, d\mu \qquad \text{(by Case 1)}$$
$$= \int \sum_{i=1}^{k} \alpha_i \mathbb{1}_{A_i}\, d\mu$$
$$= \int f\, d\mu.$$

Case 3. $f \geq 0$.

Take a sequence of non-negative simple functions $(f_n)$

such that $\quad f_n \uparrow f$. By the Monotone convergence thm

$$E f(X) = \lim_{n\to\infty} E f_n(X)$$

$$= \lim_{n\to\infty} \int f_n \, d\mu \qquad \text{( by Case 3)}.$$

$$= \int f \, d\mu .$$

Case 4. Write $f = f^+ - f^-$.

Then $\quad E f^+(X) < \infty, \quad E f^-(X) < \infty$.

So $\quad E f(X) = E f^+(X) - E f^-(X)$

$$= \int f^+ \, d\mu - \int f^- \, d\mu$$

$$= \int f \, d\mu . \qquad \blacksquare$$

As a consequence of the above thm, we can compute the expected value of functions of R.V.'s by taking integrations on the real line.

Example : Let $X$ have an exponential distribution with parameter 1. That is, $X$ has a density

$$f(x) = \begin{cases} e^{-x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Notice that

$$d\mu(x) = e^{-x}\,dx \quad \text{for } x>0,$$

where $\mu$ is the distribution of $X$.

Hence by Thm 1.9.

$$EX^k = \int_0^\infty x^k \cdot f(x)\,dx$$

$$= \int_0^\infty x^k e^{-x}\,dx.$$

$$= k! \qquad \text{for } k = 0, 1, 2, \dots$$

$$\mathrm{Var}(X) = EX^2 - (EX)^2$$

$$= 2 - 1^2 = 1.$$

Example. Let $X$ have a Poisson distribution with parameter $\lambda > 0$, i.e.

$$P\{X = k\} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

Let $\mu$ be the distribution of $X$. Then

$$\mu = \sum_{k=0}^\infty e^{-\lambda} \cdot \frac{\lambda^k}{k!}\, \delta_{\{k\}}$$

Dirac measure at $k$

$$
\begin{aligned}
EX^2 &= \int x^2 \, d\mu(x) \\
&= \sum_{k=0}^{\infty} k^2 \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} k^2 e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} k \, e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
&= \sum_{k=1}^{\infty} (k-1) e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
&\quad + \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
&= \sum_{k=2}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{(k-2)!} + \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
&= \lambda^2 + \lambda .
\end{aligned}
$$

$$
\begin{aligned}
EX &= \sum_{k=0}^{\infty} k \, e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
&= \lambda .
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= EX^2 - (EX)^2 \\
&= (\lambda^2 + \lambda) - \lambda^2 = \lambda .
\end{aligned}
$$

## §1.6 Product measures and Fubini Thm.

- Let $(X, \mathcal{A}, \mu_1)$ and $(Y, \mathcal{B}, \mu_2)$ be two $\sigma$-finite measure spaces.

- Consider the product space $(X \times Y, \mathcal{A} \times \mathcal{B})$, where $\mathcal{A} \times \mathcal{B}$ is the $\sigma$-algebra generated by the collection

$$\mathcal{T} := \{ A \times B : A \in \mathcal{A}, B \in \mathcal{B} \}.$$

Each element in $\mathcal{T}$ is called a rectangle.

- There is a unique measure $\mu$ on $(X \times Y, \mathcal{A} \times \mathcal{B})$ such that

$$\bar{\mu}(A \times B) = \mu_1(A)\, \mu_2(B) \quad \text{for all } A \in \mathcal{A},\ B \in \mathcal{B}.$$

The measure $\mu$ is often denoted as $\mu_1 \times \mu_2$.

Thm (Fubini Thm). Let $f : X \times Y \to \mathbb{R}$ be $\mathcal{A} \times \mathcal{B}$-measurable. Suppose $f \geq 0$, or $\int |f| \, d\mu < \infty$. Then

$$\int_Y \int_X f(x,y) \, d\mu_1(x) \, d\mu_2(y) = \int_{X \times Y} f \, d\bar{\mu}$$
$$= \int_X \int_Y f(x,y) \, d\mu_2(y) \, d\mu_1(x).$$

# Chap 2. Law of large numbers.

## § 2.1 Independence.

Let $(\Omega, \mathcal{F}, P)$ be a prob. space. Any element $A \in \mathcal{F}$ is called an event.

**Def.** Two events $A, B$ are called independent if

$$P(A \cap B) = P(A)\, P(B).$$

**Def.** Two r.v.'s $X$ and $Y$ are independent if

$$P\{X \in A,\ Y \in B\} = P\{X \in A\}\, P\{Y \in B\} \quad \text{for any } A, B \in \mathcal{B}(\mathbb{R}).$$

i.e. $\{X \in A\}$ and $\{Y \in B\}$ are independent for all $A, B \in \mathcal{B}(\mathbb{R})$.

**Def.** Two $\sigma$-algebras $\mathcal{F}$ and $\mathcal{G}$ are independent if for all $A \in \mathcal{F}$, $B \in \mathcal{G}$, the events $A$ and $B$ are independent.

**Remark**:

- Two r.v.'s $X$ and $Y$ are independent
  $\Leftrightarrow$ $\sigma(X)$ and $\sigma(Y)$ are independent.

  Recall that $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\}$.

- Two events $A, B$ are independent
  $\Leftrightarrow$ $\mathbb{1}_A$ and $\mathbb{1}_B$ are independent.

Def. • $\sigma$-algebras $\mathcal{F}_1, \cdots, \mathcal{F}_n$ are independent if

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i) \quad \text{for all } A_i \in \mathcal{F}_i,\ i=1,2,\cdots,n.$$

• r.v.'s $X_1, \cdots, X_n$ are independent if

$$P\left(\bigcap_{i=1}^{n} X_i^{-1}(A_i)\right) = \prod_{n=1}^{\infty} P\{X_i \in A_i\} \quad \text{for all } A_i \in \mathcal{B}(\mathbb{R}),$$
$$i=1,\cdots,n.$$

• Events $A_1, \cdots, A_n$ are independent if for any $I \subset \{1,2,\cdots,n\}$

$$P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i).$$

Equivalently, $A_1, \cdots, A_n$ are independent if

$$\mathbb{1}_{A_1}, \cdots, \mathbb{1}_{A_n} \text{ are independent.}$$

Next we give some sufficient conditions for independence.

Def. • Collections of sets $\mathcal{A}_1, \cdots, \mathcal{A}_n \subset \mathcal{F}$ are independent if

$$P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i) \quad \text{for any } A_i \in \mathcal{A}_i \text{ and } I \subset \{1,\cdots,n\}.$$

• Assume $\Omega \in \mathcal{A}_i$ for all $1 \le i \le n$. Then

$$\mathcal{A}_1, \cdots, \mathcal{A}_n \text{ are independent} \iff P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

for all $A_i \in \mathcal{A}_i$, $i=1,\cdots,n$.

Def. ① A collection $\mathcal{A}$ of subsets of $\Omega$ is said to be a $\pi$-system if

$$A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}.$$

② A collection $\mathcal{L}$ of subsets of $\Omega$ is said to be a $\lambda$-system if

(i) $\Omega \in \mathcal{L}$ ; (ii) If $A, B \in \mathcal{L}$, $A \subset B$, then $B \setminus A \in \mathcal{L}$;
(iii) If $A_n \in \mathcal{L}$ and $A_n \uparrow A$, then $A \in \mathcal{L}$.

Thm ( Dynkin's $\pi$-$\lambda$ Thm )

Suppose that $\mathcal{P}$ is a $\pi$-system and $\mathcal{L}$ is a $\lambda$-system such that

$$\mathcal{P} \subset \mathcal{L}.$$

Then $\sigma(\mathcal{P}) \subset \mathcal{L}$.

Pf. Let $\ell(\mathcal{P})$ be the smallest $\lambda$-system that contains $\mathcal{P}$.
We will show that $\ell(\mathcal{P})$ is a $\sigma$-algebra. This implies that

$$\sigma(\mathcal{P}) \subset \ell(\mathcal{P}) \subset \mathcal{L}.$$

We divide the proof of $\ell(\mathcal{P})$ being a $\sigma$-algebra into several steps.

Step 1 : A $\lambda$-system that is closed under intersection is a $\sigma$-algebra.

Check: ① $A \in \mathcal{L}$, $\Omega \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$.

② $A_n \in \mathcal{L}$, $n = 1, 2, \cdots$

$$\Rightarrow A_n^c \in \mathcal{L}, \quad \text{and} \quad \bigcap_{k=1}^{n} A_n^c \in \mathcal{L}$$

Moreover $\bigcup_{k=1}^{n} A_n = \left( \bigcap_{k=1}^{n} A_n^c \right)^c \in \mathcal{L}$.

Since $\bigcup_{k=1}^{n} A_n \uparrow \bigcup_{k=1}^{\infty} A_n \Rightarrow \bigcup_{k=1}^{\infty} A_n \in \mathcal{L}$.

Step 2: $\ell(\mathcal{P})$ is closed under intersection.

Set $\mathcal{G}_A = \{ B : A \cap B \in \ell(\mathcal{P}) \}$. We claim that

- If $A \in \ell(\mathcal{P})$, then $\mathcal{G}_A$ is a $\lambda$-system.

To prove this claim, we note that

(1) $\Omega \in \mathcal{G}_A$;

(2) If $B, C \in \mathcal{G}_A$, $B \subset C$, then

$$(C \setminus B) \cap A = (C \cap A) \setminus (B \cap A)$$
$$\in \ell(\mathcal{P})$$

So $C \setminus B \in \mathcal{G}_A$.

(3) If $B_n \in \mathcal{G}_A$, $B_n \uparrow B$, then

$$A \cap B_n \in \ell(\mathcal{P}) \quad \text{and} \quad A \cap B_n \uparrow A \cap B$$

which implies $A \cap B \in \ell(\mathcal{P})$, $\Rightarrow B \in \mathcal{G}_A$.

To see that $\ell(\mathcal{P})$ is closed under intersection, notice that $\mathcal{P}$ is a $\pi$-system. So

$$\left.\begin{array}{l} \text{If } A \in \mathcal{P}, \text{ then } \mathcal{G}_A \supset \mathcal{P} \\ \text{but } \mathcal{G}_A \text{ is a } \lambda\text{-system} \end{array}\right\} \Rightarrow \mathcal{G}_A \supset \ell(\mathcal{P})$$

Thus if $A \in \mathcal{P}$ and $B \in \ell(\mathcal{P}) \Rightarrow A \cap B \in \ell(\mathcal{P})$.

which implies $A \in \ell(\mathcal{P}),\ B \in \mathcal{P} \Rightarrow A \cap B \in \ell(\mathcal{P})$.

Hence $\left.\begin{array}{l} A \in \ell(\mathcal{P}) \Rightarrow \mathcal{G}_A \supset \mathcal{P} \\ \mathcal{G}_A \text{ is a } \lambda\text{-system} \end{array}\right\} \Rightarrow \mathcal{G}_A \supset \ell(\mathcal{P})$

Hence $\ell(\mathcal{P})$ is closed under intersection.

$\blacksquare$

As an application of Dynkin's $\pi$-$\lambda$ Thm, we have.

Thm 2.1. If $\mathcal{A}_1, \cdots, \mathcal{A}_n$ are independent and each $\mathcal{A}_i$ is a $\pi$-system, then

$$\sigma(\mathcal{A}_1), \cdots, \sigma(\mathcal{A}_n) \text{ are independent.}$$

Pf. We first show that $\sigma(\mathcal{A}_1), \mathcal{A}_2, \cdots, \mathcal{A}_n$ are independent.

Let $A_i \in \mathcal{A}_i$, $i = 2, 3, \cdots, n$. Write $F = A_2 \cap \cdots \cap A_n$.

Set

$$\mathcal{L} = \{ A : P(A \cap F) = P(A) P(F) \}.$$

Then ① $\mathcal{L}$ is a $\lambda$-system. (by def of $\lambda$-system)

② $\mathcal{L} \supset \mathcal{A}_1$. (by the independence of $\mathcal{A}_1, \cdots, \mathcal{A}_n$)

Since $\mathcal{A}_1$ is a $\pi$-system, so by Dynkin's $\pi$-$\lambda$ Thm, $\mathcal{L} \supset \sigma(\mathcal{A}_1)$.

This implies that for each $A \in \sigma(\mathcal{A}_1)$, $A_i \in \mathcal{A}_i$, $i = 2, \cdots, n$,

$$P(A \cap A_2 \cap \cdots \cap A_n) = P(A) \cdot P\left(\bigcap_{i=2}^{n} A_i\right) = P(A) \cdot P(A_2) \cdots P(A_n).$$

Hence $\sigma(A_1), \mathcal{A}_2, \cdots, \mathcal{A}_n$ are independent. That is,

$\mathcal{A}_2, \mathcal{A}_3, \cdots, \mathcal{A}_n, \sigma(\mathcal{A}_1)$ are independent.

Since $\mathcal{A}_2$ is a $\pi$-system, the previous argument

$\Rightarrow \sigma(\mathcal{A}_2), \mathcal{A}_3, \cdots, \mathcal{A}_n, \sigma(\mathcal{A}_1)$ are independent.

An iterated argument shows that $\sigma(\mathcal{A}_1), \cdots, \sigma(\mathcal{A}_n)$ are independent. $\blacksquare$.