2.

Imitating the conventional expectation, we define:

Def

$\mathscr{D} = \{D_1, \cdots, D_k\}$: decomposition of Ω.

$\xi = \sum_{j=1}^{\ell} x_j I_{A_j}(\omega)$, where $A_j = \{\xi = x_j\}$, a random variable on Ω.

Define $\boxed{E(\xi, \mathscr{D})} = \sum_{j=1}^{\ell} x_j P(A_j \mid \mathscr{D})$, called the $\boxed{\text{conditional expectation of } \xi \text{ w.r.t. } \mathscr{D}}$.

Sometimes we denote $E(\xi, \mathscr{D})$ by $\boxed{E(\xi,\mathscr{D})(\omega)}$ to emphasize it to be a random variable.

Rmk:

① Recalling that $P(A_j \mid \mathscr{D}) = \sum_{i=1}^{k} P(A_j \mid D_i) I_{D_i}(\omega)$.

Thus, $E(\xi,\mathscr{D}) = \sum_{j=1}^{\ell} x_j \left(\sum_{i=1}^{k} P(A_j \mid D_i) I_{D_i}(\omega) \right) = \sum_{i=1}^{k} \left(\sum_{j=1}^{\ell} x_j P(A_j \mid D_i) \right) I_{D_i}(\omega)$.

In other words, $E(\xi, \mathscr{D})$ is the random variable w/ value $\sum_{j=1}^{\ell} x_j P(A_j \mid D_i)$ on each D_i.

Prop

(a) $E(a\xi + b\eta \mid \mathscr{D}) = a E(\xi \mid \mathscr{D}) + b E(\eta \mid \mathscr{D})$, a, b : consts.

(b) $E(\xi \mid \Omega) = E\xi$, the const. random variable.

(c) $E(C \mid \mathscr{D}) = C$, where C : a const. random variable.

(d) If $\xi = I_A(\omega)$, then $E(\xi \mid \mathscr{D}) = P(A \mid \mathscr{D})$.

(e) $E E(\xi \mid \mathscr{D}) = E\xi$. (formula for total probability)

② (Easier way to memorize)

Define $\boxed{E(\xi \mid D_i)} = \sum_{j=1}^{\ell} x_j P(A_j \mid D_i)$, the conditional expectation of ξ w.r.t. D_i,

$E(\xi \mid \mathscr{D}) = \sum_{i=1}^{k} E(\xi \mid D_i) I_{D_i}$

Def

$\mathscr{D} = \{D_1, \cdots, D_k\}$: decomposition of Ω.

η : random variable on Ω.

We say η is $\boxed{\text{measurable w.r.t. } \mathscr{D}}$ (or $\boxed{\mathscr{D}\text{-measurable}}$) if $\mathscr{D}_\eta \preccurlyeq \mathscr{D}$. (i.e. $\mathscr{D}$: finer than $\mathscr{D}_\eta$).

Rmk:

η : $\mathscr{D}$-measurable $\iff$ η can be represented as $\eta(\omega) = \sum_{i=1}^{k} y_i I_{D_i}(\omega)$, where some y_i might equal.

$\iff$ η takes const. values on atoms of $\mathscr{D}$.

Example:

① If $\mathscr{D} = \{\Omega\}$, then η : $\mathscr{D}$-measurable $\Leftrightarrow \eta \equiv C$, a const. random variable.

② Every random variable η is $\mathscr{D}_\eta$-measurable.

Prop

η : $\mathscr{D}$-measurable

$\Rightarrow E(\xi\eta|\mathscr{D}) = \eta E(\xi|\mathscr{D})$.

In particular, $E(\eta|\mathscr{D}) = \eta$ (and $E(\eta|\mathscr{D}_\eta) = \eta$.)

⟨Pf⟩

Check the equality on each D_i. #

Prop

$\mathscr{D}_1, \mathscr{D}_2$: decompositions. w/ $\mathscr{D}_1 \preceq \mathscr{D}_2$. ξ : random variable.

Then
$$E[E(\xi|\mathscr{D}_2)|\mathscr{D}_1] = E(\xi|\mathscr{D}_1).$$

⟨Pf⟩

Check directly. #

Def

$\xi, \eta_1, \cdots, \eta_k$: random variables.

$\mathscr{D}_{\eta_1, \cdots, \eta_k}$: decomposition induced by $\eta_1, \cdots, \eta_k$.

$\boxed{E(\xi|\eta_1, \cdots, \eta_k)} := E(\xi|\mathscr{D}_{\eta_1, \cdots, \eta_k})$, called the **conditional expectation of ξ w.r.t. $\eta_1, \cdots, \eta_k$.**

Prop

(a) If ξ and η are indep., then $E(\xi|\eta) = E\xi$, the const. random variable.

(b) $E(\eta|\eta) = \eta$.

(c)
$$E[E(\xi|\eta_1, \eta_2)|\eta_1] = E(\xi|\eta_1).$$

(d) If ξ and η are independent and identically distributed random variables, then $E(\xi|\xi+\eta) = E(\eta|\xi+\eta) = \frac{\xi+\eta}{2}$.

⟨Pf⟩

(a) ~ (c) is standard check!

For (d), see P82, Example 4. #

The deduction is easy!!

Chap II. Mathematical Foundations of Probability Theory.

§1. Probabilistic Model for an Experiment w/ Infinitely Many Outcomes. Kolmogorov's Axioms.

1.

Idea: In constructing probabilistic models for uncountable spaces Ω, we must assign probabilities, not to individual outcomes, but to subsets of Ω.

Def **algebra** (containing Ω, closed under finite union and complement.)

Def $\mathscr{A}$: algebra on Ω.

$\mu : \mathscr{A} \to [0, \infty]$ is called a **finitely additive measure** if $\mu(A+B) = \mu(A) + \mu(B)$, $\forall$ disjoint pair $A, B \in \mathscr{A}$. μ: **finite** if $\mu(\Omega) < \infty$. μ: **finitely additive probability measure** or **finitely additive probability** if $\mu(\Omega) = 1$.

2.

Def An ordered triple $(\Omega, \mathscr{A}, P)$, where Ω: set, $\mathscr{A}$: algebra on Ω, P: finitely additive probability is called a **probabilistic model in the extended sense**.

Rmk: The above concept is too broad to have a fruitful mathematical theory. Hence, we will need the following:

Def **σ-algebra**.

Def **measurable space**. (set + σ-algebra).

Def μ : finitely additive measure on $(\Omega, \mathscr{A})$. algebra.

μ: **countably additive** (or **σ-additive**) or simply a **measure** if $\forall$ pairwise disjoint $A_1, A_2, \cdots \in \mathscr{A}$ w/ $\sum A_n \in \mathscr{A}$, we have $\mu(\sum A_n) = \sum \mu(A_n)$.

μ: **σ-finite** if $\Omega = \sum_{n=1}^{\infty} \Omega_n$, where $\Omega_n \in \mathscr{A}$, w/ $\mu(\Omega_n) < \infty$, $\forall n$.

Def P: countably additive measure on $(\Omega, \mathscr{A})$.

P is called a **probability measure** or simply a **probability** if $P(\Omega) = 1$.

The following theorem gives conditions under which a finitely additive measure is countably additive.

Thm.

P: finitely additive measure on $(\Omega, \mathscr{A})$, w/ $P(\Omega) = 1$.

Then T.F.A.E.

(1) P: σ-additive

(2) P: cont. from below

(3) P: cont. from above.

(4) P: cont. at $\emptyset$.

3.

Def (Kolmogorov's axiom system).
An ordered triple $(\Omega, \mathcal{F}, P)$, where
(1) Ω : set (2) $\mathcal{F}$: σ-algebra on Ω (3) P : probability on $\mathcal{F}$, is called
a $\boxed{\text{probabilistic model}}$ or a $\boxed{\text{probability space}}$.
Ω : $\boxed{\text{sample space}}$ or $\boxed{\text{space of elementary events}}$
a set A in $\mathcal{F}$: $\boxed{\text{event}}$.
$P(A)$: $\boxed{\text{probability of the event } A}$.

$\boxed{\text{End of Week 6}}$