ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

MODELOS ESTADÍSTICOS PARA LA DETECCIÓN DE PATRONES EN MEDIO AMBIENTE Y FINANZAS

ESTUDIO DE ALGUNOS ASPECTOS DE LA TEORÍA DE PROBABILIDADES A TRAVÉS DE LA TEORÍA DE CATEGORÍAS

TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERO MATEMÁTICO

WILLIAM GABRIEL GRANDA BETANCOURT

william.granda@epn.edu.ec

DIRECTOR: PH. D. MIGUEL ALFONSO FLORES SÁNCHEZ

miguel.flores@epn.edu.ec

DMQ, FEBRERO 2022

CERTIFICACIONES

Yo, WILLIAM GABRIEL GRANDA BETANCOURT, declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

William Gabriel Granda Betancourt

Certifico que el presente trabajo de integración curricular fue desarrollado por William Gabriel Granda Betancourt, bajo mi supervisión.

Ph. D. Miguel Alfonso Flores Sánchez
DIRECTOR

DECLARACIÓN DE AUTORÍA

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el(los) producto(s) resultante(s) del mismo, es(son) público(s) y estará(n) a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

William Gabriel Granda Betancourt

Ph. D. Miguel Alfonso Flores Sánchez

RESUMEN

En el presente trabajo se estudian algunos aspectos de la teoría de probabilidades a través de la teoría de categorías. Para ello, se comienza con un repaso de teoría de la medida y teoría de probabilidades. Luego, se desarrollan los conceptos básicos de teoría de categorías y para ilustrarlos se presentan ejemplos de los mismos en el contexto probabilístico. Finalmente, se presenta la categoría **Prob** de espacios de probabilidad, en la cual se reformulan las componentes centrales de la teoría de la probabilidad; como la medibilidad, la independencia y la esperanza condicional, sobre las construcciones teóricas de categorías.

Palabras clave: teoría de categorías, teoría de probabilidades, categoría **Prob**.

ABSTRACT

In the present work some aspects of probability theory are studied through category theory. To do this, it begins with a review of measure theory and probability theory. Then, the basic concepts of category theory are developed and examples of them are presented in the probabilistic context to illustrate them. Finally, the **Prob** category of probability spaces is presented, in which the central components of probability theory are reformulated; such as measurability, independence and conditional expectation, on the theoretical constructions of categories.

Keywords: category theory, probability theory, **Prob** category.

Índice general

Índice de figuras

Capítulo 1

Descripción del componente desarrollado

La estructura de este proyecto es la siguiente:

En primer lugar, se presentan algunos resultados de teoría de la medida y los conceptos generales de teoría de probabilidades. También, se desarrolla la definición de esperanza condicional. Este capítulo termina con una generalización del teorema de Radon-Nikodym para medidas con signo.

Luego, se presentan los conceptos básicos de teoría de categorías y para ilustrarlos se exponen ejemplos empleando herramientas de teoría de la medida y teoría de probabilidades, la mayoría de estos ejemplos provienen de los trabajos de Franz [2], Giry [3], Lawrence [6] y Perrone [10]. Por ejemplo, de [3] exponemos el funtor de Giry, el cual se define usando herramientas de teoría de la medida. Por otro lado, de Franz [2] exponemos la categoría **mpProb**, en la cual se introduce una noción de independencia de flechas en **mpProb**. Este capítulo finaliza con la mónada de Giry.

Finalmente, gracias a las herramientas mencionadas anteriormente se estudia la categoría **Prob**; definida en [13], cuyos objetos son los espacios de probabilidad y cuyas flechas entre espacios de probabilidad son las funciones que preservan conjuntos de medida nula. Sobre esta categoría se expone la generalización de algunas herramientas clásicas de la teoría de la probabilidad. Por ejemplo, se define el funtor esperanza condicional, el cual corresponde a una generalización de la esperanza condicional usual y contrastamos las propiedades que satisface este funtor con las propiedades que satisface su contraparte en teoría de probabilidades. Una vez definido este funtor se da una generalización de la medibilidad, este concepto depende de una flecha en **Prob**. Para terminar esta sección presentamos una noción de independencia en **Prob** y algunos resultados para estudiar la relación entre el funtor esperanza condicional y la independencia en **Prob**.

1.1. Objetivo general

Estudiar la categoría **Prob** de espacios de probabilidad, y las generalizaciones de la esperanza condicional, la medibilidad y la independencia, usando herramientas y métodos de la teoría de categorías.

1.2. Objetivos específicos

1.3. Objetivos específicos

1. Revisar los trabajos de Franz [2], Giry [3] y Lawvere [6], en los cuales se encuentra la base de la aplicación de la teoría de categorías a la teoría de probabilidades.

2. Revisar [10] para proponer ejemplos de los conceptos de teoría de categorías en teoría de probabilidades.

3. Estudiar los conceptos básicos sobre teoría de categorías, para de esta forma presentar una generalización de la esperanza condicional, la independencia y la medibilidad.

4. Usar el teorema de Radon-Nikodym para demostrar la existencia del funtor esperanza condicional en Prob, el cual es una generalización de la esperanza condicional usual.

5. Demostrar que el funtor esperanza condicional es una generalización de la esperanza condicional usual.

1.4. Alcance

Para alcanzar nuestro objetivo es necesario adquirir un conocimiento básico en teoría de categorías y un conocimiento más profundo de la teoría de probabilidades.

Por ejemplo, haremos un estudio de los conceptos básicos de la teoría de categorías, como la definición de categoría, funtor, transformación natural. Además, revisaremos los trabajos de Franz [2], Giry [3] y Lawrence [6], los cuales nos servirán para desarrollar los conceptos de categorías empleando herramientas de teoría de la medida y teoría de probabilidades. Otros ejemplos son tomados de [10] y [13] y desarrollados a detalle.

1.5. Marco teórico

La “probabilidad categórica” es una colección de estructuras y métodos categóricos que se aplican a la teoría de probabilidad, la teoría de la medida y la estadística matemática. El campo de la probabilidad categórica tiene como objetivo reformular los componentes centrales de la teoría de la probabilidad sobre las construcciones teóricas de categorías. Esto incluye herramientas como la composición de mapas probabilísticos, las distribuciones de probabilidad marginales y conjuntas, la independencia, las probabilidades condicionales, entre otras [3],[6].

- **Justificación Teórica:**

  Los ejemplos más destacados de la aplicación de la teoría de categorías a la teoría de probabilidades son los enfoques de Giry [3] y Lawvere [6]. Sin embargo, existen pocas categorías con objetos como espacios de probabilidad debido a la dificultad de encontrar una condición apropiada para definir flechas entre ellos.

  En [6] Lawvere construyó lo que denominó la categoría de “probabilistic mapping”, denotada por **Stoch**, en la cual se presentan las primeras ideas sobre cómo utilizar la teoría de categorías para extender funciones a “flechas estocásticas”. Precisamente, esta es la idea para definir las flechas en la categoría **Prob**, se consideran las funciones medibles y se las extiende a flechas en **Prob** imponiendo una condición de continuidad sobre las mismas.

Por su parte, Giry en [3] presenta una definición categórica de los procesos aleatorios y se proporcionan herramientas para su estudio. Un concepto central en la mayoría de los enfoques categóricos de la teoría de probabilidad es el de una "mónada de probabilidad", introducido por primera vez por Giry en [3]. Por otro lado, en [5] se puede encontrar una lista detallada con la mayoría de las mónadas de probabilidad en la literatura, junto con sus principales propiedades.

Un trabajo reciente con flechas apropiadas lo realizan en [8], se introduce una noción de "flechas acotadas" entre los espacios de probabilidad y se define la categoría de todos los espacios de probabilidad y todas las flechas acotadas entre los mismos, esta categoría se denota por **CPS**.

En [12] se proporciona una categoría simple denotada por χ, en la cual se formula otra generalización de la esperanza condicional y se presenta un enfoque categórico para abordar la teoría de riesgo. La generalización de la esperanza condicional se presenta como un functor de la categoría opuesta de χ, denotada por $\chi^{\mathbf{op}}$, a la categoría **Set**. No obstante, los objetos y flechas de esta categoría son limitados, y por esta razón, no se pueden usar como base de una teoría de probabilidades.

- **Justificación Metodológica**:

  La teoría de la probabilidades no se trata solo de espacios de medidas de probabilidad, sino también, y sobre todo de las interacciones y propagaciones de diferentes variables aleatorias. Esto puede tratarse categóricamente en términos de categorías monoidales y functores monoidales [9]. Fritz y Perrone abordan las distribuciones de probabilidad conjunta y marginal en términos de estructuras monoidales y muestran cómo obtener una "categoría de espacios de probabilidad" a partir de una categoría equipada con una mónada de probabilidad.

  En [13] se presenta la categoría **Prob**, la cual tiene por objetos a los espacios de probabilidad y como flechas a las funciones medibles que conservan conjuntos de medida nula entre ellos. Se prueba que $\mathbf{CPS^{op}}$ es una sub-categoría de **Prob**.

Otro resultado importante de [13] es que una vez definida la categoría **Prob** se prueba la existencia del functor de esperanza condicional, el cual es una generalización de la noción clásica de esperanza condicional. Se presenta también la generalización de algunas nociones importantes relacionadas con una σ-álgebra en la teoría de probabilidad clásica, como lo son, la independencia y la medibilidad, estos dos conceptos dependen de una flecha en **Prob**.

- **Justificación práctica:**

  La condición impuesta sobre las flechas en la categoría **Prob** es adecuada y esto permite generalizar algunas herramientas clásicas en la teoría de la probabilidad. Así, esta categoría sirve como base para establecer una teoría de probabilidades. Por ejemplo, haciendo uso de la categoría **Prob**, en [12] se presenta un modelo binomial de fijación de precios de activos en un entorno categórico. Al aplicar las herramientas de teoría de categorías a los modelos binomiales de fijación de precios se obtienen resultados más generales, como se puede ver en [12].

1.5.1. Teoría de la Medida

En esta sección vamos a presentar la definición de medidas y medidas con signo. Además, se enuncia el teorema de descomposición de Jordan, el cual nos permite encontrar una representación para una medida con signo como una diferencia de dos medidas no negativas. Finalmente, introducimos el espacio medible producto y el teorema de Fubini.

Conocer los conceptos básicos de teoría de la medida es necesario para revisar los trabajos de Giry [3], Lawvere [6] y Franz [2], con lo cual, podremos proponer ejemplos de los conceptos básicos de teoría de categorías empleando herramientas de teoría de la medida y nos servirá para estudiar las propiedades que satisface el funtor esperanza condicional.

La demostración de los teoremas y resultados se puede revisar en [4] o [15].

Medidas y medidas con signo

Una medida con signo generaliza el concepto de medida permitiendo que una medida esté definida sobre una $\sigma-$álgebra a valores en $\overline{\mathbb{R}} = [-\infty, +\infty]$. A lo largo de esta sección consideraremos a $(X, \mathcal{F})$ como un espacio medible, es decir, X es un conjunto no vacío y $\mathcal{F}$ una $\sigma-$álgebra definida sobre X.

En primer lugar, recordemos la definición de una medida:

Definición 1.1 (Medida, espacio medido). *Sea $(X, \mathcal{F})$ un espacio medible. Una función $\mu\colon \mathcal{F} \to \overline{\mathbb{R}}_+$ se dice una* medida *sobre $(X, \mathcal{F})$ si:*

- $m(\emptyset) = 0$, *y*
- *Para toda colección $\{F_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ de conjuntos disjuntos dos a dos, tal que $\bigcup\limits_{n\in\mathbb{N}} F_n \in \mathcal{F}$, se tiene que*

$$m\left(\bigcup_{n\in\mathbb{N}} F_n\right) = \sum_{n\in\mathbb{N}} m(F_n).$$

La tripleta $(X, \mathcal{F}, \mu)$ se denomina espacio medido *o* espacio de medida*.*

Definición 1.2 (Medida con signo). *Sean $(X, \mathcal{F})$ un espacio medible y $\mu\colon \mathcal{F} \to \overline{\mathbb{R}}$ una función de conjuntos. Se dice que μ es una* medida con signo *si:*

- $\mu(\emptyset) = 0$;
- *Para toda colección $\{F_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ de conjuntos disjuntos dos a dos, se tiene que*
$$\mu\left(\bigcup_{n\in\mathbb{N}} F_n\right) = \sum_{n\in\mathbb{N}} \mu(F_n).$$
- *μ asume como máximo uno de los valores $-\infty$ o $+\infty$.*

Además, se dice que la medida con signo es finita *si no toma los valores $-\infty$ o $+\infty$.*

OBSERVACIÓN. Se tiene que toda medida es una medida con signo. Pero una medida con signo, en general, no es una medida.

Teoremas de descomposición de medidas

El objetivo de esta sección es presentar que toda medida con signo μ se puede escribir como la diferencia $\mu = \mu^+ - \mu^-$, donde μ^+ y μ^- son medidas.

Previo a exponer el teorema de la descomposición de Jordan, consideremos las siguientes definiciones con respecto a una medida con signo:

DEFINICIÓN 1.3 (Parte positiva y parte negativa, variación). *Sean $(X, \mathcal{F})$ un espacio medible y μ una medida con signo definida sobre $(X, \mathcal{F})$.*

- *Para $x \in X$, se definen la* parte positiva *de μ por $\mu^+(x) = \text{máx}\{\mu(x), 0\}$ y la* parte negativa *de μ por $\mu^-(x) = \text{máx}\{-\mu(x), 0\}$.*
- *La* variación *de μ es la medida (o* medida positiva*), denotada por $|\mu|$ y definida por*
$$|\mu| = \mu^+ + \mu^-.$$

DEFINICIÓN 1.4 (Medidas con signo mutuamente singulares, continuidad absoluta). *Sean $(X, \mathcal{F})$ un espacio medible y μ, ν dos medidas con signo definidas sobre $(X, \mathcal{F})$.*

- *Se dice que μ y ν son* mutuamente singulares*, denotado por $\mu \perp \nu$, si existen dos conjuntos disjuntos F, G de X, tales que $X = F \cup G$, y para todo conjunto medible E, se satisface*
$$|\mu|(E \cap F) = |\nu|(E \cap G) = 0.$$
- *Se dice que ν es* absolutamente continua *con respecto a μ, denotado por $\nu \ll \mu$, si*
$$|\mu|(A) = 0 \implies |\nu|(A) = 0,$$
para todo $F \in \mathcal{F}$.

Así, presentamos el teorema de descomposición de Jordan, el cual será de utilidad para el desarrollo de este trabajo.

Teorema 1.1 (Teorema de descomposición de Jordan). *Sean $(X, \mathcal{F})$ un espacio medible y μ una medida con signo definida sobre $(X, \mathcal{F})$. Existen dos medidas mutuamente singulares, denotadas por μ^+ y μ^-, y definidas sobre $(X, \mathcal{F})$ tales que*

$$\mu = \mu^+ - \mu^-.$$

Además, el par (μ^+, μ^-) es único.

La demostración del teorema de descomposición de Jordan se puede encontrar en [4] o en el Anexo A.1.

Observación.

- El par (μ^+, μ^-) se conoce como la descomposición de Jordan de la medida con signo μ.
- El teorema de descomposición de Jordan nos permite generalizar ciertos resultados para las medidas con signo, análogos a los resultados que se tienen para las medidas.

Finalmente, consideremos la siguiente proposición, la cual relaciona la continuidad absoluta de una medida con signo con la continuidad absoluta de su parte positiva y su parte negativa.

Proposición 1.2. *Sean $(X, \mathcal{F})$ un espacio medible y μ, ν medidas con signo definidas sobre $(X, \mathcal{F})$. Si $\mu \ll \nu$, entonces $\mu^+ \ll \nu$ y $\mu^- \ll \nu$.*

Demostración. Anexo 1, Demostración 1.2. □

Ahora, desarrollamos de manera sucinta la integración con respecto a las medidas con signo.

Definición 1.5 (Función simple). *Sea $(X, \mathcal{F}, \mu)$ un espacio medido con μ una medida con signo. Se dice que una función $f\colon X \to \mathbb{R}$ es una* función simple *si existen $\{A_k\}_{k=1}^n \subseteq \mathcal{F}$ una colección de conjuntos disjuntos y $\{a_k\}_{k=1}^n \subseteq \mathbb{R}$ tales que*

$$f(x) = \sum_{k=1}^{n} a_k \mathbb{1}_{A_k}.$$

Además, definimos su integral con respecto a la medida con signo μ, como *sigue*

$$\begin{aligned}\int_X f\,d\mu &= \sum_{k=1}^{n} a_k \mu(A_k),\\ &= \sum_{k=1}^{n} a_k \mu^+(A_k) - \sum_{k=1}^{n} a_k \mu^+(A_k),\\ &= \int_X f\,d\mu^+ - \int_X f\,d\mu^-.\end{aligned}$$

OBSERVACIÓN. Gracias a la descomposición de la medida con signo $\mu = \mu^+ - \mu^-$ se puede considerar la integral de una función simple con respecto a una medida con signo μ como la diferencia de dos integrales usuales.

La observación precedente nos motiva a presentar la siguiente definición:

DEFINICIÓN 1.6 (Integral con respecto a una medida con signo)**.** *Sean $(X, \mathcal{F}, \mu)$ un espacio medido con μ una medida con signo, y $f\colon X \to \mathbb{R}$ una función $\mathcal{F}/\mathcal{B}(\mathbb{R})-$ medible. Se define la integral de la función f con respecto a la medida con signo μ, como sigue*

$$\int_X f\,d\mu = \int_X f\,d\mu^+ - \int_X f\,d\mu^-.$$

OBSERVACIÓN. Como se mencionó, el hecho de descomponer la integral de una función medible con respecto a una medida con signo resulta de gran utilidad; no obstante, algunos resultados no se cumplen para estas integrales. Por ejemplo, si $f, g\colon X \to \overline{\mathbb{R}}_+$ son dos funciones medibles tales que

$$f \leq g,$$

en general, no se cumple que $\displaystyle\int_X f\,d\mu \leq \int_X g\,d\mu$.

1.5.2. Teoría de Probabilidades

En esta sección desarrollaremos ciertos conceptos básicos de teoría de probabilidad. Primero, presentamos la definición de esperanza. Luego, exponemos la definición de esperanza condicional y sus propiedades. También, exponemos el teorema de Radon-Nikodym para medidas positivas. Este teorema nos permite bajo algunas hipótesis, escribir una medida como una integral con respecto a otra. Así, exhibimos una extensión del teorema de Radon-Nikodym para medidas con signo y su demostración, este teorema es clave para el desarrollo de la Sección 2.1.

De igual forma, este capítulo es importante para presentar ejemplos en un contexto probabilístico de los conceptos de teoría de categorías y para definir la categoría **Prob**.

La demostración de los resultados se encuentra en [11] o [15].

Esperanza

En esta sección consideraremos a $(X, \mathcal{F}, \mathbb{P})$ como un espacio de probabilidad.

Definición 1.7 (Esperanza)**.** *Sea η una variable aleatoria. Se define la esperanza de η, denotada por $\mathbb{E}^{\mathbb{P}}(\eta) = \int_X \eta \, d\mathbb{P}$, como la integral de η con respecto a la medida de probabilidad $\mathbb{P}$.*

Ahora, presentaremos algunas propiedades que satisface la esperanza de una variable aleatoria.

Proposición 1.3. *Sean η una variable aleatoria y $(\eta_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias. Se tienen los siguiente resultados:*

- ***Convergencia Monótona:** Si $(\eta_n)_{n\in\mathbb{N}}$ es creciente y $\lim\limits_{n\to+\infty} \eta_n = \eta$ casi seguramente, entonces*

$$\lim_{n\to+\infty} \mathbb{E}^{\mathbb{P}}(\eta_n) = \mathbb{E}^{\mathbb{P}}(\eta),$$

  y $\left(\mathbb{E}^{\mathbb{P}}(\eta_n)\right)_{n\in\mathbb{N}}$ es una sucesión creciente.

- ***Convergencia Dominada:** Si $\lim\limits_{n\to+\infty} \eta_n = \eta$ casi seguramente y existe una variable aleatoria $\xi \in \mathcal{L}(X, \mathcal{F}, \mathbb{P})$ tal que $|\eta_n| \leq \xi$ para cada $n \in \mathbb{N}$, entonces*

$$\lim_{n\to+\infty} \mathbb{E}^{\mathbb{P}_X}(|\eta_n - \eta|) = 0,$$

  de donde,

$$\lim_{n\to+\infty} \mathbb{E}^{\mathbb{P}_X}(\eta_n) = \mathbb{E}^{\mathbb{P}_X}(\eta).$$

Esperanza condicional

La esperanza condicional constituye el punto de partida de la teoría de la probabilidad moderna. Por ejemplo, es fundamental en la definición de los procesos de Markov, martingalas y en general en los procesos estocásticos.

En esta sección consideraremos a $\mathcal{L}^2(X, \mathcal{F}, \mathbb{P})$ como el espacio vectorial de las variables aleatorias real valuadas v, tales que $\int_X |v|^2 \, d\mathbb{P} < +\infty$. También, $L^2(X, \mathcal{F}, \mathbb{P})$ representará al espacio cociente de $\mathcal{L}^2((X, \mathcal{F}, \mathbb{P}))/ \sim_{\mathbb{P}_X}$, donde

$$\mu_1 \sim_{\mathbb{P}_x} \mu_2 \text{ si } \mathbb{P}\left(\{\omega \,|\, \mu_1(\omega) \neq \mu_2(\omega)\}\right) = 0.$$

El siguiente teorema nos asegura la existencia de la esperanza condicional:

TEOREMA 1.4. *Sean $(X, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $\lambda \in \mathcal{L}^1(X, \mathcal{F}, \mathbb{P})$ una variable aleatoria integrable. Sea $\mathcal{G}$ una sub$-\sigma-$álgebra de $\mathcal{F}$, existe una variable aleatoria $\eta \in \mathcal{L}^1(X, \mathcal{F}, \mathbb{P})$ tal que*

a) η es $\mathcal{G}-$medible;

b) Para toda $G \in \mathcal{G}$, se tiene que

$$\int_G \eta \, d\mathbb{P}_X = \int_G \lambda \, d\mathbb{P}_X.$$

Si η_1 es otra variable aleatoria que satisface las propiedades anteriores, entonces $\eta_1 = \eta$ casi seguramente.

La variable aleatoria η se denomina una versión de la esperanza condicional de λ dado $\mathcal{G}$ y se denota por $\mathbb{E}[\lambda|\mathcal{G}]$. Escribiremos $\eta = \mathbb{E}[\lambda|\mathcal{G}]$ casi seguramente.

OBSERVACIÓN. Para observar el significado intuitivo de la esperanza condicional, supongamos que se ha llevado a cabo un experimento y que la única información disponible sobre que punto de la muestra $\omega \in X$ ha sido elegido es el conjunto de valores $Z(\omega)$ para cada variable aleatoria $\mathcal{G}-$medible. Se tiene que $Y(\omega) = \mathbb{E}[\lambda|\mathcal{G}](\omega)$ es el valor esperado de $\lambda(\omega)$ dada esta información.

La forma más común de demostrar el teorema 1.4 es mediante el teorema de Radon-Nikodym, como se puede ver en [11]. Sin embargo, en [15] se presenta una demostración de la existencia de $\mathbb{E}[\lambda|\mathcal{G}]$ en $\mathcal{L}^2(X, \mathcal{F}_X, \mathbb{P}_X)$ usando el hecho de que el espacio $L^2(X, \mathcal{F}_X, \mathbb{P}_X)$ es completo. Luego, se prueba su existencia en $\mathcal{L}^1(X, \mathcal{F}_X, \mathbb{P}_X)$.

Para terminar esta sección, consideremos las propiedades que satisface la esperanza condicional, las cuales serán contrastadas con las propiedades que satisface la generalización de la esperanza condicional definida en la Categoría **Prob**.

PROPOSICIÓN 1.5. *Sean $\lambda \in \mathcal{L}(X, \mathcal{F}, \mathbb{P})$ y $\mathcal{G}_1, \mathcal{G}_2$ dos sub-$\sigma-$álgebras de $\mathcal{F}$. Se tienen los siguiente resultados:*

1. *Si η es una versión de $\mathbb{E}[\lambda|\mathcal{G}_1]$, entonces $\mathbb{E}[\eta] = \mathbb{E}[\lambda]$.*
2. ***Linealidad:** Sean $\lambda_1, \lambda_2 \in \mathcal{L}(X, \mathcal{F}, \mathbb{P})$ y $\alpha, \beta \in \mathbb{R}$, se tiene que*
$$\mathbb{E}[\alpha\lambda_1 + \beta\lambda_2|\mathcal{G}_1] = \alpha\mathbb{E}[\lambda_1|\mathcal{G}_1] + \beta\mathbb{E}[\lambda_2|\mathcal{G}_1].$$
3. ***Positividad:** Si $\lambda \geq 0$, entonces $\mathbb{E}[\lambda|\mathcal{G}_1] \geq 0$.*
4. ***Monotonía:** Sea $\eta \in \mathcal{L}(X, \mathcal{F}_X, \mathbb{P}_X)$, si $\lambda \leq \eta$, entonces $\mathbb{E}[\xi|\mathcal{G}] \leq \mathbb{E}[\eta|\mathcal{G}]$ casi seguramente.*
5. ***Convergencia Monótona:** Si $(\xi_n)_{n\in\mathbb{N}}$ es una sucesión creciente de variables aleatorias no negativas en $\mathcal{L}(X, \mathcal{F}, \mu)$ tales que $\lim\limits_{n\to+\infty} \lambda_n = \xi$, entonces $(\mathbb{E}[\lambda_n|\mathcal{G}_1])_{n\in\mathbb{N}}$ es una sucesión creciente y*
$$\lim_{n\to+\infty} \mathbb{E}[\lambda_n|\mathcal{G}_1] = \mathbb{E}[\lambda|\mathcal{G}_1],$$
casi seguramente.
6. *Si η es una variable aleatoria $\mathcal{G}_1-$medible y acotada, entonces*
$$\mathbb{E}[\eta\lambda|\mathcal{G}_1] = \eta\mathbb{E}[\lambda|\mathcal{G}_1],$$
casi seguramente.
7. *Si λ es independiente de $\mathcal{G}_1$, entonces $\mathbb{E}[\lambda|\mathcal{G}_1] = \mathbb{E}[\lambda]$.*

Teorema de Radon-Nikodym

Sean $(X, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $\mathbb{Q}$ una medida de probabilidad absolutamente continua con respecto a $\mathbb{P}$. Si $\lambda \in \mathcal{L}(X, \mathcal{F}, \mathbb{P})$ es tal que

$$\mathbb{Q}(F) = \int_F \lambda \, d\mathbb{P},$$

para cada $F \in \mathcal{F}$, se dice que λ es la derivada de Radon-Nikodym de la medida de probabilidad $\mathbb{Q}$ en $(X, \mathcal{F})$, emplearemos la notación $\lambda = \frac{d\mathbb{Q}}{d\mathbb{P}}$ en $(X, \mathcal{F})$.

El teorema de Radon-Nikodym nos asegura la existencia y la unicidad casi segura de la derivada de Radon-Nikodym de una medida de probabilidad.

Teorema 1.6 (Teorema de Radon-Nikodym)**.** *Sean $(X, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $\mathbb{Q}$ una medida $\sigma-$finita sobre $(X, \mathcal{F})$ y absolutamente continua con respecto a $\mathbb{P}$. Entonces existe λ una variable aleatoria no negativa en $\mathcal{L}(X, \mathcal{F}, \mathbb{P})$ tal que*

$$\mathbb{Q}(F) = \int_X \lambda \, \mathbb{P},$$

para cada $F \in \mathcal{F}$. La variable aleatoria λ es única casi seguramente.

Notemos que en el teorema 1.6 suponemos que $(X, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad. Sin embargo, podemos reescribir el mismo en términos de medidas $\sigma-$finitas como sigue:

Teorema 1.7 (Teorema de Radon - Nikodym)**.** *Sean $(X, \mathcal{F}, \mu)$ un espacio medido $\sigma-$finito y ν una medida $\sigma-$finita sobre $(X, \mathcal{F})$. Si ν es absolutamente continua con respecto a μ, entonces existe una función integrable $f \in \mathcal{L}(X, \mathcal{F}, \mu)$ tal que*

$$\nu(F) = \int_F f \, d\mu,$$

para todo $F \in \mathcal{F}$. La función f es única casi en todas partes. Se dice que f es la derivada de Radon-Nikodym y se la denota por $\frac{d\nu}{d\mu}$.

Ahora, consideremos una extensión del teorema de Radon-Nikodym para medidas con signo.

COROLARIO 1.8. *Sean $(X, \mathcal{F}, \mu)$ un espacio medido $\sigma-$finito y ν una medida con signo $\sigma-$finita y absolutamente continua con respecto a μ. Existe una función medible f tal que*

$$\nu(F) = \int_F f \, d\mu,$$

para todo $F \in \mathcal{F}$. La función f es única casi en todas partes. Se dice que f es la derivada de Radon-Nikodym y se la denota por $\frac{d\nu}{d\mu}$.

Demostración. Anexo A.4, Demostración A.4. □

Expondremos algunas propiedades con respecto a la derivada de una medida con signo, que serán empleadas para demostrar las propiedades que satisface la generalización de la esperanza condicional presentada en el Sección 2.1.

PROPOSICIÓN 1.9. *Sean $(X, \mathcal{F}, \mu)$ un espacio medido finito y ν una medida con signo definida sobre $(X, \mathcal{F})$ absolutamente continua con respecto a μ. Además, supongamos que la derivada de Radon-Nikodym, denotada por $\frac{d\nu}{d\mu}$ de la medida ν con respecto a la medida μ existe. Se tienen los siguientes resultados:*

- *Si $\nu = \nu^+ - \nu^-$ es la descomposición de Jordan de la medida con signo ν, entonces las derivadas de Radon-Nikodym $\frac{d\nu^+}{d\mu}$ y $\frac{d\nu^-}{d\mu}$ de las medidas ν^+ y ν^-, respectivamente, existen.*

  Además, si (P, N) es una descomposición de Hahn de X con respecto a la medida ν, entonces

  $$\frac{d\nu^+}{d\mu} = \mathbb{1}_P \frac{d\nu}{d\mu} \quad \text{y} \quad \frac{d\nu^+}{d\mu} = -\mathbb{1}_N \frac{d\nu}{d\mu},$$

  casi todas partes.

- *Si $\left(\frac{d\nu}{d\mu}\right)^+$ y $\left(\frac{d\nu}{d\mu}\right)^-$ representan la parte positiva y la parte negativa de la derivada de Radon-Nikodym de ν, respectivamente, entonces*

  $$\frac{d\nu^+}{d\mu} = \left(\frac{d\nu}{d\mu}\right)^+ \quad \text{y} \quad \frac{d\nu^-}{d\mu} = \left(\frac{d\nu}{d\mu}\right)^+,$$

  casi todas partes.

Demostración. Anexo A.4, Demostración A.4. $\square$

Capítulo 2

Metodología

2.1. Teoría de Categorías

En esta sección vamos a desarrollar los conceptos básicos de teoría de categorías y para ilustrarlos presentaremos ejemplos empleando herramientas de teoría de la medida y teoría de probabilidades.

Además, emplearemos esta sección para presentar los resultados fundamentales de Franz [2], Giry [3] y Lawvere [6] como ejemplos de los conceptos generales de teoría de categorías. También, se presentan ejemplos de [10], en este artículo se pueden encontrar los conceptos de teoría categorías con ejemplos en diferentes áreas de las matemáticas, como: teoría de grupos, teoría de grafos, probabilidad, entre otros.

Específicamente, de [2] introducimos la categoría **mpProb**, sobre la cual se define una noción de "independencia" y se prueba que corresponde a una extensión de la independencia clásica en teoría de probabilidades. Además, mostramos que esta categoría no posee productos. Mientras que de Giry [3] presentamos el funtor de Giry y la *mónada de Giry*, la cual se estudia a detalle en la Sección 2.1.5. Por su parte de [6], exponemos la categoría de flechas estocásticas y demostramos que satisface las propiedades que definen una categoría.

2.1.1. Categorías

Definición 2.1 (Categoría)**.** *Una* categoría $\mathcal{C}$ *consiste de:*

- *Una colección de objetos, denotada por* $\textbf{Ob}(\mathcal{C})$*;*
- *Una colección de flechas, denotada por* $\boldsymbol{\mathcal{A}}(\mathcal{C})$*;*
- *Cada flecha tiene asignada dos operaciones, denominadas dominio y codominio. El dominio asigna a cada flecha f un objeto $A = \text{dom}(f)$, y el codominio asigna a cada flecha f un objeto $B = \text{cod}(f)$. Emplearemos la notación*
$$f\colon A \to B.$$
- *Para cada par de flechas $f, g \in \boldsymbol{\mathcal{A}}(C)$, tales que $\text{dom}(g) = \text{cod}(f)$, se define la operación composición de f y g, denotada por $g \circ f$, tal que $g \circ f\colon \text{dom}(f) \to \text{cod}(g)$. De ser así, diremos que f y g son flechas componibles.*
- *Sean $f, g, h \in \boldsymbol{\mathcal{A}}(C)$ flechas tales que $\text{dom}(g) = \text{cod}(f)$ y $\text{dom}(h) = \text{cod}(g)$, se tiene que*
$$h \circ (g \circ f) = (h \circ g) \circ f,$$
es decir, la composición es asociativa.
- *Para cada objeto $A \in \textbf{Ob}(\mathcal{A})$, existe una flecha $1_A\colon A \to A$, a la que llamamos flecha identidad, tal que si $f\colon A \to B$ y $g\colon C \to A$ son flechas, entonces se satisface que*
$$f \circ 1_A = f \quad \text{y} \quad g \circ 1_A = g.$$

Observación. Si A, B son dos objetos de $\mathcal{C}$, la colección de flechas $f\colon A \to B$ la denotaremos por $\boldsymbol{\mathcal{C}}(A, B)$.

Una de las herramientas más importantes en la teoría de categorías es el razonamiento en términos de diagramas. Presentamos una definición informal, pero consistente, de un diagrama. Una definición formal de una diagrama en términos de funtores se la puede encontrar en [7].

DEFINICIÓN 2.2. *Sea $\mathcal{C}$ una categoría, un* diagrama *en $\mathcal{C}$ es un multidígrafo formado a partir de objetos y flechas de $\mathcal{C}$, tales que se satisfacen las siguientes propiedades:*

- *Cada objeto y cada flecha puede aparecer más de una vez en el diagrama;*
- *Entre dos objetos cualesquiera puede existir también más de una flecha;*
- *Para cada objeto, la flecha identidad está implícitamente presente en el diagrama, pero generalmente no se la dibuja;*
- *Para cada par de flechas $f\colon A \to B$ y $g\colon B \to C$, su composición $g \circ f$ está implícitamente presente en el diagrama, pero generalmente no se dibuja.*

DEFINICIÓN 2.3. *Se dice que un diagrama es* conmutativo *(o* conmuta*) si para cada par de objetos $A, B \in$ **Ob**$(\mathcal{C})$ en el diagrama, todas las composiciones a lo largo de caminos de flechas componibles que conectan a A y B son iguales.*

OBSERVACIÓN. No todos lo diagramas son conmutativos.

EJEMPLO 1. El diagrama

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "h"'] & B \arrow[d, "g"] \\ C \arrow[r, "k"'] & D \end{tikzcd}$$

conmuta si y solo si $g \circ f = k \circ h$.

Para ilustrar el concepto de categoría, primero presentamos algunos ejemplos clásicos de categorías.

EJEMPLO 2. 1. **Categoría uno:** Esta categoría posee un objeto y una flecha, la cual corresponde a la flecha identidad de su único objeto.

$$\bullet$$

La denotaremos por **1**.

2. **Categoría dos:** Es la categoría que tiene dos objetos, una flecha identidad para cada objeto, y una única flecha entre los dos objetos. Su diagrama es:

   $$\bullet \to \bullet$$

   Se denota por **2**.

3. **Categoría tres:** Esta categoría tiene tres objetos, tres flechas de identidad para cada objeto, exactamente una flecha del primero al segundo objeto, exactamente una flecha del segundo al tercer objeto, y exactamente una flecha del primero al tercero objeto. Su diagrama viene dado por:

   $$\begin{tikzcd} \bullet \arrow[r] \arrow[dr] & \bullet \arrow[d] \\ & \bullet \end{tikzcd}$$

   Esta categoría será denotada por **3**.

4. **Categoría de conjuntos:** Es la categoría definida por:

   - **Objetos**: Conjuntos;
   - **Flechas**: Funciones entre conjuntos;
   - **Composición**: Composición usual de funciones;
   - **Identidad:** Función identidad definida sobre un conjunto.

5. **Preórdenes:** Sean X un conjunto y $\leq$ una relación sobre X, $\leq$ se dice un preorden si es reflexiva y transitiva. Se define la categoría **Pre** por medio de la siguiente información: Sus objetos son los elementos de X, dados dos objetos $x, y \in X$ existe una única flecha de x hacia y si y solo si $x \leq y$, la composición entre dos flechas está dada por la transitividad y las identidades para cada objeto están dadas por la reflexividad.

Dado que la composición de funciones medibles es medible podemos definir la categoría de los espacios medibles, presentada en [3]:

EJEMPLO 3. Categoría de los espacios medibles: Se define la categoría **Meas** cuyos objetos son espacios medibles y cuyas flechas entre espacios medibles son las funciones medibles.

OBSERVACIÓN. Notemos que la composición en la categoría **Meas** es la composición usual de funciones y la flecha identidad en **Meas** es la función identidad definida en un espacio medible $(X, \mathcal{F}_X)$, la cual es $\mathcal{F}_X-$medible.

Ahora, presentamos la categoría χ, definida en [1]. La categoría χ es la base para abordar la teoría de riesgo desde un enfoque categórico.

DEFINICIÓN 2.4 (Categoría χ). *Sean $(X, \mathcal{F}_X)$ un espacio medible y $\chi := \chi(X, \mathcal{F}_X)$ el conjunto de todos los pares de la forma $(\mathcal{G}, \mathbb{P}_X)$, donde $\mathcal{G} \subset \mathcal{F}_X$ es una sub$-\sigma-$álgebra de $\mathcal{F}_X$ y $\mathbb{P}_X$ es una medida de probabilidad definida sobre $(X, \mathcal{F}_X)$. Para un elemento $\mathcal{U} \in \chi$, denotamos su $\sigma-$álgebra y su medida de probabilidad por $\mathcal{F}_\mathcal{U}$ y $\mathbb{P}_\mathcal{U}$, respectivamente, es decir, $\mathcal{U} = (\mathcal{F}_\mathcal{U}, \mathbb{P}_\mathcal{U})$. Para $\mathcal{U}, \mathcal{V} \in \chi$, consideremos la relación binaria, denotada por $\leq_\chi$, y definida como sigue*

$$\mathcal{V} \leq_\chi \mathcal{U} \text{ si y solo si } \mathcal{F}_\mathcal{V} \subset \mathcal{F}_\mathcal{U} \text{ y } \mathbb{P}_\mathcal{V} \gg \mathbb{P}_\mathcal{U}.$$

Se define la categoría χ como aquella que tiene exactamente una flecha $f_\mathcal{U}^\mathcal{V} \colon \mathcal{V} \to \mathcal{U}$ en χ si y solo si $\mathcal{V} \leq_\chi \mathcal{U}$.

OBSERVACIÓN. Notemos que $(\chi, \leq_\chi)$ es un conjunto preordenado. En efecto, sea $\mathcal{U} \in \chi$, la reflexividad se cumple dado que $\mathcal{F}_\mathcal{U} \subseteq \mathcal{F}_\mathcal{U}$ y dado que $\mathbb{P}_\mathcal{U}$ es absolutamente continua con respecto a sí misma. Así, resta probar la transitividad, para ello, sean $\mathcal{U}, \mathcal{V}, \mathcal{W} \in \chi$ tales que

$$\mathcal{V} \leq_\chi \mathcal{U} \quad \text{y} \quad \mathcal{U} \leq_\chi \mathcal{W},$$

vamos a demostrar que $\mathcal{V} \leq_\chi \mathcal{W}$. Por un lado, tenemos que

$$\mathcal{F}_\mathcal{V} \subset \mathcal{F}_\mathcal{U} \quad \text{y} \quad \mathcal{F}_\mathcal{U} \subset \mathcal{F}_\mathcal{W},$$

de donde, $\mathcal{F}_\mathcal{V} \subset \mathcal{F}_\mathcal{W}$.

Por otro lado, sea $A \in \mathcal{F}_{\mathcal{V}}$ tal que $\mathbb{P}_{\mathcal{V}}(A) = 0$, debemos probar que $\mathbb{P}_{\mathcal{W}}(A) = 0$. Dado que $\mathbb{P}_{\mathcal{U}} \ll \mathbb{P}_{\mathcal{V}}$, tenemos que $\mathbb{P}_{\mathcal{U}}(A) = 0$, de donde, usando el hecho de que $\mathbb{P}_{\mathcal{W}} \ll \mathbb{P}_{\mathcal{U}}$, obtenemos que $\mathbb{P}_{\mathcal{W}}(A) = 0$.

Por lo tanto, la categoría χ está bien definida. Retomaremos está categoría posteriormente para presentar ejemplos de otros conceptos de teoría de categorías.

Para exponer el siguiente ejemplo, es necesario considerar la siguiente definición:

Definición 2.5 (Función acotada (Motoyama-Tanaka [8])). *Sean $(X, \mathcal{F}_X, \mathbb{P}_X)$ y $(Y, \mathcal{F}_Y, \mathbb{P}_Y)$ dos espacios de probabilidad. Sea $f\colon (Y, \mathcal{F}_Y, \mathbb{P}_Y) \to (X, \mathcal{F}_X, \mathbb{P}_X)$ una función $\mathcal{F}_Y/\mathcal{F}_X-$ medible, se dice que f es* acotada *si existe un número positivo $M > 0$ tal que*

$$\mathbb{P}_Y(f^{-1}(A)) \leq M\mathbb{P}_X(A),$$

para cada $A \in \mathcal{F}_X$.

Así, tenemos el siguiente ejemplo de una categoría en el contexto probabilístico:

Definición 2.6 (Categoría **CPS**). *La* categoría **CPS** *se define por medio de la siguiente información:*

- ***Objetos:*** *Espacios de probabilidad;*
- ***Flechas:*** *Sean $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ y $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ dos espacios de probabilidad, el conjunto de flechas entre estos espacios se define por*

  $$\textbf{CPS}(\overline{X}, \overline{Y}) = \{f | f\colon \overline{X} \to \overline{Y} \text{ es una función acotada}\}.$$
- ***Composición:*** *Composición usual de funciones;*
- ***Identidades:*** *Si $Id\colon (X, \mathcal{F}_X) \to (X, \mathcal{F}_X)$ representa la función identidad (la misma que es $\mathcal{F}_X-$medible), entonces la flecha identidad para un espacio de probabilidad en **CPS** está denotada por*

  $$id\colon (X, \mathcal{F}_X, \mathbb{P}_X) \to (X, \mathcal{F}_X, \mathbb{P}_X).$$

Notemos que id es una función acotada. En efecto, para todo $A \in \mathcal{F}_X$, se tiene que

$$\mathbb{P}_X\left(id^{-1}(A)\right) = \mathbb{P}_X(A).$$

OBSERVACIÓN. CPS es una categoría.

Previo a demostrar que **CPS** es una categoría consideremos el siguiente resultado con respecto a la imagen inversa de la composición de funciones: Sean $f\colon X \to Y$ y $G\colon Y \to Z$ funciones, si $C \subseteq Z$, entonces $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.

En efecto, sea $x \in (g \circ f)^{-1}(C)$, se tienen las siguientes equivalencias:

$$\begin{aligned}(g \circ f)(x) \in C &\iff g(f(x)) \in C,\\ &\iff f(x) \in g^{-1}(C),\\ &\iff x \in f^{-1}(g^{-1}(C)),\end{aligned}$$

es decir, hemos probado que

$$(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)). \tag{2.1}$$

Ahora, para demostrar que **CPS** está bien definida basta con verificar que la composición de funciones acotadas es acotada, para ello, sean $f\colon (X, \mathcal{F}_X, \mathbb{P}_X) \to (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ y $g\colon (Y, \mathcal{F}_Y, \mathbb{P}_Y) \to (Z, \mathcal{F}_Z, \mathbb{P}_Z)$ dos flechas en **CPS**, vamos a demostrar que $g \circ f\colon (X, \mathcal{F}_X, \mathbb{P}_X) \to (Z, \mathcal{F}_Z, \mathbb{P}_Z)$ es acotada. Sea $A \in \mathcal{F}_Z$, debemos probar que existe $M > 0$ tal que

$$\mathbb{P}_X\left((g \circ f)^{-1}(A)\right) \leq M\mathbb{P}_Z(A).$$

En primer lugar, sabemos existe $M_1 > 0$ tal que

$$\mathbb{P}_Y(g^{-1}(A)) \leq M_1\mathbb{P}_Z(A). \tag{2.2}$$

De igual forma, existe $M_2 > 0$ tal que

$$\mathbb{P}_X\left(f^{-1}\left(g^{-1}(A)\right)\right) \leq M_2\mathbb{P}_Y(g^{-1}(A), \tag{2.3}$$

para $g^{-1}(A) \in \mathcal{F}_Y$.

Luego, combinando (2.2) y (2.3), obtenemos que

$$\mathbb{P}_X\left(f^{-1}\left(g^{-1}(A)\right)\right) \leq M_2 \cdot M_1 \mathbb{P}_Z(A),$$

de donde, tomando $M := M_1 \cdot M_2 > 0$, se sigue el resultado.

Por lo tanto, hemos probado que $g \circ f$ es una flecha en **CPS**.

Observación. De hecho, en [8] se define la categoría **CMS**, cuyos objetos son los espacios medidos y las flechas entre dos espacios medidos $(X, \mathcal{F}_X, \mu)$ y $(Y, \mathcal{F}_Y, \nu)$ son las funciones acotadas en el sentido de la definición 2.5.

Presentamos también la categoría **mpProb** definida en Franz [2], para esto, consideremos la siguiente definición:

Definición 2.7. *Sean $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$, $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ espacios de probabilidad y $f\colon \overline{Y} \to \overline{X}$ una función $\mathcal{F}_Y/\mathcal{F}_X-$medible, decimos que f* preserva medidas *si $\mathbb{P}_Y \circ f^{-1} = \mathbb{P}_X$.*

Definición 2.8 (Categoría **mpProb**). *La categoría **mpProb** tiene por objetos a los espacios de probabilidad y como flechas a las funciones que preservan medidas en el sentido de la definición 2.7.*

Observación. La identidad de la definición 2.7 es demasiado estricta para el tratamiento categórico de las flechas, por ello, en [8] se prueba que **CPS** es una extensión de la categoría **mpProb**.

Proposición 2.1. ***mpProb** es una categoría.*

Demostración. Es fácil ver que las flechas de **mpProb** satisfacen la ley de asociatividad y las leyes de identidades. Así, basta con demostrar que la composición de funciones que preservan medidas, también preserva medidas. Para ello, sean $f\colon \overline{Y} \to \overline{X}$ y $g^{-}\colon \overline{Z} \to \overline{Y}$ funciones que preservan medidas, es decir,

$$\mathbb{P}_X = \mathbb{P}_Y \circ f^{-1} \quad \text{y} \quad \mathbb{P}_Y = \mathbb{P}_Z \circ g^{-1}. \tag{2.4}$$

Vamos a demostrar que $f \circ g\colon \overline{Z} \to \overline{X}$ preserva medidas, es decir, debemos probar que

$$\mathbb{P}_X = \mathbb{P}_Z \circ (f \circ g)^{-1}.$$

Junto con (2.4) y la definición de medida imagen de $\mathbb{P}_Y$ con respecto a la función f, tenemos que

$$\mathbb{P}_Y \circ f^{-1} = (\mathbb{P}_Z \circ g^{-1}) \circ f^{-1},$$

de donde, nuevamente, usando (2.4), se sigue que

$$\mathbb{P}_X = (\mathbb{P}_Z \circ g^{-1}) \circ f^{-1} = \mathbb{P}_Z \circ (f \circ g)^{-1},$$

como queríamos. □

Ahora, presentamos la categoría de funciones estocásticas definida en [6], para esto, es necesario considerar la siguiente definición:

DEFINICIÓN 2.9 (Función estocástica)**.** *Sean* $(X, \mathcal{F}_X)$ *y* $(Y, \mathcal{F}_Y)$ *espacios medibles. Una* función estocástica $T\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ *es una función*

$$T\colon X \times \mathcal{F}_Y \to [0, 1],$$

tal que

1. *Para cada* $x \in X$*, la función* $T_x := T(x, \cdot)\colon \mathcal{F}_Y \to [0, 1]$ *es una medida de probabilidad en* $(Y, \mathcal{F}_Y)$*, es decir,* T_x *satisface las siguientes propiedades:*
   a) $0 \leq T_x(F_Y) \leq 1$ *para cada* $F_Y \in \mathcal{F}_Y$*;*
   b) $T_x(Y) = 1$*;*
   c) $T_x\left(\bigcup_{n\in\mathbb{N}} F_n\right) = \sum_{n\in\mathbb{N}} T_x(F_n)$ *para cualquier colección* $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{F}_Y$ *de eventos disjuntos dos a dos.*
2. *Para cada* $F_Y \in \mathcal{F}_Y$*, la función* $T_{F_Y} := T(\cdot, F_Y)\colon X \to [0, 1]$ *es* $\mathcal{F}_X/\mathcal{B}([0, 1])-$*medible*

Llamaremos a $T(x, F_Y)$ *la* T*-probabilidad (condicional) del evento* F_Y *en* $(Y, \mathcal{F}_Y)$ *dado el punto* x *en* $(X, \mathcal{F}_X)$.

Presentemos algunos ejemplos para ilustrar el concepto de función estocástica:

Ejemplo 4 (Función característica)**.** Sea $(X, \mathcal{F}_X)$ un espacio medible, se define la función característica $\delta\colon X \times \mathcal{F}_X \to \mathbb{R}$ por

$$\delta(x, F_X) = \begin{cases} 1 & \textbf{si} \quad x \in F_X, \\ 0 & \textbf{si} \quad x \notin F_X. \end{cases}$$

para cada $x \in X$ y para cada $F_X \in \mathcal{F}_X$.

Notemos que la función $\delta_x = \delta(x, \cdot)$ corresponde a la medida de Dirac en el punto $x \in X$.

Por otro lado, es fácil ver que, para cada $F_X \in \mathcal{F}_X$, $\delta(\cdot, F_X)$ es una función $\mathcal{F}_X/\mathcal{B}(\mathbb{R})-$ medible.

Ejemplo 5 (Funciones estocásticas deterministas)**.** Sean $(X, \mathcal{F}_X), (Y, \mathcal{F}_Y)$ espacios medibles y $f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ una función $\mathcal{F}_X/\mathcal{F}_Y-$medible. Se define la función estocástica $\delta_f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ dada por

$$\begin{aligned} \delta_f \colon & X \times \mathcal{F}_Y \longrightarrow [0,1] \\ & (x, F_Y) \longmapsto \delta_f(x, F_Y) = \begin{cases} 1 & \textbf{si} \quad f(x) \in F_Y, \\ 0 & \textbf{si} \quad f(x) \in F_Y^C. \end{cases} \end{aligned}$$

Llamaremos a δ_f la función estocástica determinista inducida por f.

Definición 2.10 (Composición de funciones estocásticas)**.** *Sean* $T\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ *y* $U\colon (Y, \mathcal{F}_Y) \to (Z, \mathcal{F}_Z)$ *dos funciones estocásticas. Se define la composición* $U \circ T\colon (X, \mathcal{F}_X) \to (Z, \mathcal{F}_Z)$ *de* U *y* T *por la integral*

$$U \circ T(x, F_Z) = \int_Y U(\cdot, F_Z)\, dT_x = \int_{y \in Y} U(y, F_Z)\, T(x, dy).$$

para cada $x \in X$ *y* $F_Z \in \mathcal{F}_Z$*.*

Ejemplo 6. Sean $f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ y $g\colon (Y, \mathcal{F}_Y) \to (Z, \mathcal{F}_Z)$ funciones medibles. Para ejemplificar la composición de dos funciones estocásticas consideremos δ_f y δ_g las funciones estocásticas deterministas inducidas por f y g, respectivamente. Sean $x \in X$ y $F_Z \in \mathcal{F}_Z$, tenemos que

$$\delta_g \circ \delta_f(x, F_Z) = \int_Y \delta_g(\cdot, F_Z)\, d\delta_{f,x}. \tag{2.5}$$

Por otro lado, notemos que

$$\delta_g(x, F_Z) = \begin{cases} 1 & \text{si} \quad g(x) \in F_Z, \\ 0 & \text{si} \quad g(x) \in F_Z^C. \end{cases} = \begin{cases} 1 & \text{si} \quad x \in g^{-1}(F_Z), \\ 0 & \text{si} \quad x \in g^{-1}\left(F_Z^C\right). \end{cases}$$

Así, junto con (2.5), se tiene que

$$\begin{aligned} \delta_g \circ \delta_f(x, F_Z) &= \int_{g^{-1}(F_Z)} \delta_g(\cdot, F_Z)\, d\delta_{f,x} + \int_{g^{-1}(F_Z^C)} \delta_g(\cdot, F_Z)\, d\delta_{f,x}, \\ &= \int_{g^{-1}(F_Z)} d\delta_{f,x}, \\ &= \delta_f(x, g^{-1}(F_Z)), \end{aligned}$$

de donde, se tiene que

$$\delta_g \circ \delta_f(x, F_Z) = \begin{cases} 1 & \text{si} \quad f(x) \in g^{-1}(F_Z), \\ 0 & \text{si} \quad f(x) \in g^{-1}(F_Z^C). \end{cases} = \begin{cases} 1 & \text{si} \quad g(f(x)) \in F_Z, \\ 0 & \text{si} \quad g(f(x)) \in F_Z^C. \end{cases}$$

es decir, hemos probado que

$$\delta_f \circ \delta_g = \delta_{g \circ f}.$$

OBSERVACIÓN. Podemos ver que la integral de (2.10) está bien definida, pues $U(\cdot, F_Z)$ es una función medible y T_x es una medida de probabilidad definida sobre $(Y, \mathcal{F}_Y)$.

Nos restar verificar que $U \circ T$ es una función estocástica.

- Sea $x \in X$, probemos que $U \circ T(x, \cdot)$ es una medida de probabilidad sobre $(Z, \mathcal{F}_Z)$. Para ello, vamos a demostrar que
    - $0 \leq (U \circ T)_x(F_Z) \leq 1$ para cada $F_Z \in \mathcal{F}_Z$.
      Sean $F_Z \in \mathcal{F}_Z$ e $y \in Y$, dado que U_y es una medida de probabilidad sobre $(Z, \mathcal{F}_Z)$, se tiene que

$$0 \leq U(y, F_Z) \leq 1,$$

de donde, por la monotonía de la integral, se sigue que

$$0 \leq U \circ T(x, F_Z) \leq 1.$$

- $U \circ T(x, Z) = 1$. En efecto, tenemos que

$$\begin{aligned} U \circ T(y, Z) &= \int_Y U(y, Z)\, T(x, dy), \\ &= \int_Y T(x, dy), \\ &= T(x, Y), \\ &= 1. \end{aligned}$$

- Sea $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{F}_Z$ una colección de eventos de $\mathcal{F}_Z$ disjuntos dos a dos, debemos probar que $U \circ T\left(x, \bigcup_{n\in\mathbb{N}} F_n\right) = \sum_{n\in\mathbb{N}} U \circ T(x, F_n)$.

  Se tiene que

$$\begin{aligned} (U \circ T)\left(x, \bigcup_{n\in\mathbb{N}} F_n\right) &= \int_Y U\left(y, \bigcup_{n\in\mathbb{N}} F_n\right)\, dT_x, \\ &= \int_Y \sum_{n\in\mathbb{N}} U(y, F_n)\, T(x, dy), \end{aligned}$$

  de donde, gracias al teorema de convergencia monótona 1.3, obtenemos que

$$\begin{aligned} (U \circ T)\left(x, \bigcup_{n\in\mathbb{N}} F_n\right) &= \sum_{n\in\mathbb{N}} \int_Y U(y, F_n)\, T(x, dy), \\ &= \sum_{n\in\mathbb{N}} U \circ T(x, F_n). \end{aligned}$$

Con esto, hemos probado que $U \circ T(x, \cdot)$ es una medida de probabilidad definida sobre $(Z, \mathcal{F}_Z)$.

- Sea $F_Z \in \mathcal{F}_Z$, vamos a demostrar que $U \circ T(\cdot, F_Z)\colon X \to [0, 1]$ es $\mathcal{F}_X/\mathcal{B}([0, 1])-$ medible.

  Probemos el resultado para funciones características, para ello, sea $f\colon (Y, \mathcal{F}_Y) \to (Z, \mathcal{F}_Z)$ una función $\mathcal{F}_X/\mathcal{F}_Y-$medible, consideremos la función estocástica determinista $\delta_f\colon (Y, \mathcal{F}_Y) \to (Z, \mathcal{F}_Z)$. Para $x \in X$ y

$F_Z \in \mathcal{F}_Z$, tenemos que

$$\begin{aligned}
\delta_f \circ T(x, F_Z) &= \int_Y \delta_f(\cdot, F_Z)\, dT_x, \\
&= \int_{f^{-1}(F_Z)} \delta_f(\cdot, F_Z)\, dT_x + \int_{f^{-1}(F_Z^C)} \delta_f(\cdot, F_Z)\, dT_x, \\
&= \int_{f^{-1}(F_Z)} \delta_f(\cdot, F_Z)\, dT_x, \\
&= \int_{f^{-1}(F_Z)} dT_x, \\
&= T(x, f^{-1}(F_Z)).
\end{aligned}$$

Con esto, tenemos que

$$\delta_f \circ T(\cdot, F_Z) = T(\cdot, f^{-1}(F_Z)),$$

de donde, dado que T_{F_Y} es $\mathcal{F}_X/\mathcal{B}(\mathbb{R})-$ medible para cada $F_Y \in \mathcal{F}_Y$ se sigue que $\delta_f \circ T(\cdot, F_Z)$ también lo es.

Gracias a la linealidad de la integral el resultado previo sigue siendo cierto si U_{F_Z} es una función simple. Luego, el caso general es consecuencia del teorema de convergencia monótona 1.3 y el hecho de que U_{F_Z} puede verse como el límite de una sucesión creciente de funciones simples.

Ejemplo 7 (Categoría **Stoch**). La categoría **Stoch** está definida por la siguiente información:

1. **Objetos:** Espacios medibles;

2. **Flechas:** Sean $\overline{X} = (X, \mathcal{F}_X)$ y $\overline{Y} = (Y, \mathcal{F}_Y)$ dos espacios medibles, el conjunto de flechas entre estos espacios se define por

$$\textbf{Stoch}(\overline{X}, \overline{Y}) = \{U | U \colon \overline{X} \to \overline{Y} \text{ es una función estocástica}\}.$$

3. **Composición**: Composición de funciones estocásticas en el sentido de la definición 2.10.

4. **Identidades**: La función estocástica identidad es la función característica definida en el ejemplo 4 y la denotaremos por $Id \colon (X, \mathcal{F}_X) \to (X, \mathcal{F}_X)$.

OBSERVACIÓN. **Stoch** es una categoría. Para probar esto, gracias a los resultados previos basta con demostrar que la composición en **Stoch** es asociativa y se satisfacen los axiomas de identidad.

- **Asociatividad:**

  Sean $T\colon (X,\mathcal{F}_X) \to (Y,\mathcal{F}_Y)$, $U\colon (Y,\mathcal{F}_Y) \to (Z,\mathcal{F}_Z)$ y $V\colon (Z,\mathcal{F}_Z) \to (W,\mathcal{F}_W)$ funciones estocásticas, vamos a demostrar que

$$V \circ (U \circ T) = (V \circ U) \circ T.$$

  Para ello, sean $x \in X$ y $F_W \in \mathcal{F}_W$, debemos probar que

$$V \circ (U \circ T)(x, F_W) = (V \circ U) \circ T(x, F_W). \tag{2.6}$$

  Por un lado, para $y \in Y$ y $F_W \in \mathcal{F}_W$, tenemos que

$$V \circ U(y, F_W) = \int_Z V(\cdot, F_W)\, dU_y,$$

  de donde, se tiene que

$$\begin{aligned} (V \circ U) \circ T(x, F_W) &= \int_Y V \circ U(\cdot, F_W)\, dT_x, \\ &= \int_Y \left(\int_Z V(\cdot, F_W)\, dU_y \right) dT_x. \end{aligned} \tag{2.7}$$

  Ahora, por la parte 3 del Lema 2.4, obtenemos que

$$\int_Y \left(\int_Z V(\cdot, F_W)\, dU_y \right) dT_x = \int_Z V(\cdot, \mathcal{F}_W)\, d(U \circ T)_x,$$

  con $U \circ T(x, \mathcal{F}_Z) = \displaystyle\int_Y U(\cdot, F_Z)\, dT_x$. Así, de (2.7), se sigue que

$$(V \circ U) \circ T(x, F_W) = \int_Z V(\cdot, \mathcal{F}_W)\, d((U \circ T)_x). \tag{2.8}$$

  Por otro lado, el lado izquierdo de (2.6) es

$$V \circ (U \circ T)(x, F_W) = \int_Z V(\cdot, F_W)\, d(U \circ T)_x.$$

Junto la identidad precedente y (2.8), podemos concluir que

$$V \circ (U \circ T)(x, F_W) = (V \circ U) \circ T(x, F_W),$$

como queríamos.

- Sean $T\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ y $U\colon (Y, \mathcal{F}_Y) \to (X, \mathcal{F}_X)$ funciones estocásticas, debemos probar que

$$Id \circ U = U \quad \text{y} \quad T \circ Id = T.$$

Para esto, sean $y \in Y$ y $F_X \in \mathcal{F}_X$, vamos a demostrar que

$$Id \circ U(y, F_X) = U(y, F_X).$$

Por definición de composición aplicada a las funciones estocásticas Id y U, tenemos que

$$\begin{aligned}
Id \circ U(y, F_X) &= \int_X Id(x, F_X)\, dU_y, \\
&= \int_X \delta(x, F_X)\, dU_y, \\
&= \int_{F_X} \delta(x, F_X)\, dU_y + \int_{F_X^C} \delta(x, F_X)\, dU_y, \\
&= \int_{F_X} dU_y, \\
&= U(y, F_X).
\end{aligned}$$

Con lo cual, hemos probado que $Id \circ U = U$.

Ahora, vamos a demostrar que

$$T \circ Id = T.$$

Sean $x \in X$ y $F_Y \in \mathcal{F}_Y$, debemos probar que

$$T \circ Id(x, \mathcal{F}_X) = T(x, \mathcal{F}_X).$$

Nuevamente, por definición de composición aplicada a las funciones estocásticas T e Id, se tiene que

$$\begin{aligned} T \circ Id(x, \mathcal{F}_X) &= \int_X T(x, F_X)\, dId_x, \\ &= \int_X T(x, F_X)\, d\delta_x, \end{aligned}$$

de donde, dado que δ_x es la medida de Dirac en el punto $x \in X$, obtenemos que

$$T \circ Id(x, \mathcal{F}_X) = T(x, F_X),$$

con esto, podemos concluir que

$$T \circ Id = T.$$

Definición 2.11 (Isomorfismo). *Sean $\mathcal{C}$ una categoría y $A, B \in$ **Ob**$(\mathcal{C})$ objetos. Una flecha $f\colon A \to B$ en $\mathcal{C}$ es un* isomorfismo *si existe una flecha $g\colon B \to A$ tal que*

$$g \circ f = 1_A \quad \textit{y} \quad f \circ g = 1_B.$$

Observación. Dada la flecha $f\colon A \to B$ se tiene que la flecha $g\colon B \to A$ es única. Así, se escribe $g = f^{-1}$. En este caso decimos que A y B son isomorfos, denotado por $A \cong B$, si existe un isomorfismo entre ellos.

Ejemplo 8. Expondremos un ejemplo de isomorfismo en la categoría **CMS**, la cual fue presentada en observación 2.1.1. Sean $(X, \mathcal{F}_X, \mu)$ un espacio medido y $c > 0$, tenemos que μ es una medida definida sobre $(X, \mathcal{F}_X)$, de donde $c \cdot \mu$ también lo es. Así, consideremos la flecha $f\colon (X, \mathcal{F}_X, \mu) \to (X, \mathcal{F}_X, c \cdot \mu)$ en **CMS** definida por

$$\begin{aligned} f :(X, \mathcal{F}_X) &\longrightarrow (X, \mathcal{F}_X) \\ x &\longmapsto x \end{aligned}$$

Se tiene que f es un isomorfismo. En efecto, definamos $g\colon (X, \mathcal{F}_X, c \cdot \mu) \to (X, \mathcal{F}_X, \mu)$ por

$$\begin{aligned} g :(X, \mathcal{F}_X) &\longrightarrow (X, \mathcal{F}_X) \\ x &\longmapsto x. \end{aligned}$$

Probemos que g es una función acotada, para ello, sea $F_X \in \mathcal{F}_X$, dado que

$g^{-1}(F_X) = F_X$, se sigue que

$$c \cdot \mu(g^{-1}(F_X)) = c \cdot \mu(F_X) \leq 2c \cdot \mu(F_X),$$

con lo cual, tomando $M = 2c > 0$ se sigue el resultado.

Además, es fácil ver que $g \circ f = Id_{(X, \mathcal{F}_X, \mu)}$ y $f \circ g = Id_{(X, \mathcal{F}_X, c \cdot \mu)}$

La siguiente definición nos permite considerar nuevas categorías a partir de las presentadas anteriormente.

DEFINICIÓN 2.12 (Categoría Opuesta)**.** *Sea $\mathcal{C}$ una categoría, se define su categoría* opuesta *o* dual, *denotada por $\mathcal{C}^{\text{op}}$, mediante la siguiente información:*

- ***Objetos:* Ob**$(\mathcal{C}^{\text{op}}) :=$ ***Ob**($\mathcal{C}$),*
- ***Flechas**: Sean $A, B \in$ **Ob**$(\mathcal{C}^{\text{op}})$, la colección de flechas entre estos objetos está definida por $\mathcal{C}^{\text{op}}(B, A) = \mathcal{C}(A, B)$.*

  Es decir, $f\colon B \to A$ es una flecha en $\mathcal{C}^{\text{op}}$ si $f\colon A \to B$ es una flecha en $\mathcal{C}$. Para referirnos a una flecha en la categoría opuesta emplearemos la notación f^{op}.
- ***Composición**: Sean $A, B, C \in$ **Ob**$(\mathcal{C}^{\text{op}})$, consideremos las flechas $f^{\text{op}}\colon B \to A$ y $g^{\text{op}}\colon C \to B$, su composición está definida por*

  $$f^{\text{op}} \circ g^{\text{op}} = (g \circ f)^{\text{op}}.$$
- ***Identidades:* *Las mismas identidades definidas en $\mathcal{C}$.***

OBSERVACIÓN. Consideremos el siguiente diagrama en $\mathcal{C}$:

$$\begin{tikzcd}
\mathrm{A} \arrow[r, "f"] \arrow[dr, "g \circ f"'] & \mathrm{B} \arrow[d, "g"] \\
& \mathrm{C}
\end{tikzcd}$$

El diagrama correspondiente en $\mathcal{C}^{\text{op}}$ es

$$\begin{tikzcd}
A & B \arrow[l, "f^{\mathrm{op}}"'] \\
& C \arrow[u, "g^{\mathrm{op}}"'] \arrow[ul, "f^{\mathrm{op}} \circ g^{\mathrm{op}}"]
\end{tikzcd}$$

DEFINICIÓN 2.13 (Subcategoría)**.** *Sea $\mathcal{C}$ una categoría. Una* subcategoría *$\mathcal{D}$ de $\mathcal{C}$ consiste de una subcolección **Ob**$(\mathcal{D})$ de **Ob**$(\mathcal{C})$ junto con, para cada $A, B \in$ **Ob**$(\mathcal{D})$, una subcolección $\mathcal{D}(A, B)$ de $\mathcal{C}(A, B)$, tal que*

1. *Para cada objeto $A \in$ **Ob**$(\mathcal{D})$, la flecha identidad $\mathbf{1}_A$ está en $\mathcal{D}$, es decir, $\mathcal{D}$ es cerrado bajo identidades.*

2. *Para cada par de flechas componibles $f \in \mathcal{D}(A, B)$ y $g \in \mathcal{D}(B, C)$, se tiene que $g \circ f$ está en $\mathcal{D}$, es decir, $\mathcal{D}$ es cerrado bajo composición.*

Además, $\mathcal{D}$ se dice una subcategoría plena *si*

$$\mathcal{D}(A, B) = \mathcal{C}(A, B),$$

*para todo $A, B \in$ **Ob**$(\mathcal{D})$.*

OBSERVACIÓN. Consideremos la categoría **Set**$_1$ cuyos objetos con lo conjuntos y cuyas flechas entre conjuntos son funciones inyectivas. Es fácil ver que **Set**$_1$ es una subcategoría de **Set**.

EJEMPLO 9. La categoría **CPS** de espacios de probabilidad y funciones acotadas, definida en 2.6, es una subcategoría plena de la categoría **CMS** de espacios medidos y funciones acotadas. Es fácil ver que se cumplen las condiciones de la definición 2.13 y que **CMS**$((X, \mathcal{F}_X, \mathbb{P}_X), (Y, \mathcal{F}_Y, \mathbb{P}_Y)) =$ **CPS**$((X, \mathcal{F}_X, \mathbb{P}_X), (Y, \mathcal{F}_Y, \mathbb{P}_Y))$ para cada $(X, \mathcal{F}_X, \mathbb{P}_X)$, $(Y, \mathcal{F}_Y, \mathbb{P}_Y) \in$ **Ob**(**CPS**).

En gran parte de las categorías que hemos definido la colección de todos los objetos es demasiado grande para formar un conjunto. No se profundizará en este tema, dado que no es el objetivo de este trabajo; sin embargo, consideremos la siguiente definición:

DEFINICIÓN 2.14. *Una categoría $\mathcal{C}$ se denomina* pequeña *si tanto la colección **Ob**$(\mathcal{C})$ de los objetos de $\mathcal{C}$ y la colección $\boldsymbol{\mathcal{A}}(\mathcal{C})$ de flechas de $\mathcal{C}$ son* *conjuntos. Caso contrario $\mathcal{C}$ se denomina* grande.

Una categoría $\mathcal{C}$ es localmente pequeña *si para cada $A, B \in$ **Ob**$(\mathcal{C})$, la colección $\mathcal{C}(A, B)$ es un conjunto.*

OBSERVACIÓN. Si una categoría es pequeña, entonces es localmente pequeña.

EJEMPLO 10.

1. La Categoría **CPS** es una categoría localmente pequeña. En efecto, para cada $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$, $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y) \in$ ***Ob***(**CPS**), tenemos que

$$\mathbf{CPS}(\overline{X}, \overline{Y}) \subset Y^X,$$

   donde Y^X representa el conjunto de funciones del conjunto X al conjunto Y.

2. De igual forma, **Set** es una categoría localmente pequeña.

DEFINICIÓN 2.15. *Sean $\mathcal{C}$ una categoría e I, $T \in$ **Ob**$(\mathcal{C})$ objetos. Se dice que I es* inicial *si, para cada $A \in$ **Ob**$(\mathcal{C})$, existe exactamente una flecha $f\colon I \to A$ en $\mathcal{C}$.*

Por otro lado, T se dice terminal *si, para cada $A \in$ **Ob**$(\mathcal{C})$, existe exactamente una flecha $f\colon A \to T$ en $\mathcal{C}$.*

OBSERVACIÓN. Vemos que el conjunto vacío es un objeto de inicial de **Set**. Además, cualquier conjunto con un elemento, es decir, un singletón es un objeto terminal. En efecto, para cada $A \in$ ***Ob***(**Set**), existe una única función tal que

$$\begin{aligned} f\colon A &\longrightarrow \{*\} \\ x &\longmapsto f(x) = *. \end{aligned}$$

2.1.2. Funtores

Un funtor es una flecha entre categorías. Formalmente, tenemos la siguiente definición:

DEFINICIÓN 2.16 (Funtor)**.** *Sean $\mathcal{C}$ y $\mathcal{D}$ dos categorías, un* funtor covariante *$F\colon \mathcal{C} \to \mathcal{D}$ consiste de:*

- *Una función*
$$\begin{aligned} F\colon \boldsymbol{Ob}(\mathcal{C}) &\longrightarrow \boldsymbol{Ob}(\mathcal{D}) \\ A &\longmapsto F(A), \end{aligned}$$

- *Para cada* $A, B \in \boldsymbol{Ob}(\mathcal{C})$*, una función*
$$\begin{aligned} F\colon \mathcal{C}(A,B) &\longrightarrow \mathcal{D}(F(A),F(B)) \\ f &\longmapsto F(f). \end{aligned}$$

  tal que
$$F(g\circ f) = F(g)\circ F(f) \quad \textit{y} \quad F(\mathbf{1}_A) = \mathbf{1}_{F(A)},$$

  para cualesquiera par de flechas componibles $f, g \in \boldsymbol{\mathcal{A}}(\mathcal{C})$ *y para cualquier objeto* $A \in \boldsymbol{Ob}(\mathcal{C})$.

  Por otro lado, si para cada $A, B \in \boldsymbol{Ob}(\mathcal{C})$ *se tiene la función*
$$F\colon \mathcal{C}(A,B) \to \mathcal{D}(F(B),F(A)),$$

  tal que
$$F(g\circ f) = F(f)\circ F(g) \quad \textit{y} \quad F(\mathbf{1}_A) = \mathbf{1}_{F(A)},$$

  entonces F *se dice un* funtor contravariante.

Observación. Notemos que se puede definir un funtor contravariante como un funtor covariante en la categoría opuesta de $\mathcal{C}$. Es decir, si $F\colon \mathcal{C} \to \mathcal{D}$ es un funtor contravariante, entonces escribiremos $F\colon \mathcal{C}^{\mathrm{op}} \to \mathcal{D}$ y lo llamaremos funtor. Así, no haremos distinción entre funtores covariantes y contravariantes.

Nuevamente, primero presentaremos ejemplos clásicos de funtores.

Ejemplo 11.
- Sean $F\colon \mathcal{C} \to \mathcal{B}$ y $G\colon \mathcal{B} \to \mathcal{D}$ dos funtores, se define el funtor composición $G\circ F$ por:
  - Para objetos:
$$\begin{aligned} G\circ F\colon \boldsymbol{Ob}(\mathcal{C}) &\longrightarrow \boldsymbol{Ob}(\mathcal{D}) \\ A &\longmapsto G(F(A)). \end{aligned}$$
  - Para flechas:
$$\begin{aligned} G\circ F\colon \quad \boldsymbol{\mathcal{C}}(A,B) &\longrightarrow \boldsymbol{\mathcal{D}}(G(F(A)),G(F(B))) \\ [f\colon A\to B] &\longmapsto [G(F(f))\colon G(F(A)) \to G(F(B))]. \end{aligned}$$

- Sea $\mathcal{C}$ una categoría, se define el functor identidad, denotado por $\mathbf{1}_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$, como sigue

$$\mathbf{1}_{\mathcal{C}}(A) = A \quad \text{y} \quad \mathbf{1}_{\mathcal{C}}(f) = f,$$

para todo objeto $A \in \boldsymbol{Ob}(\mathcal{C})$ y para toda flecha $f \in \boldsymbol{\mathcal{A}}(\mathcal{C})$.

- Sean A, B dos conjuntos preordenados, un funtor covariante entre las categorías correspondientes es exactamente una función creciente.

El ejemplo precedente nos permite presentar la siguiente observación:

OBSERVACIÓN. Denotaremos por **Cat** a la categoría de todas las categorías pequeñas y por **LCat** a la categoría de todas las categorías localmente pequeñas, la composición de funtores en estas categorías se define como el funtor composición y las flechas identidad corresponden a los funtores identidad.

Por su parte, ejemplos de funtores en el contexto probabilístico serían:

EJEMPLO 12.

- Sean la categoría **CPS** y la categoría **Meas**, se define el functor de olvido U: **CPS** $\to$ **Meas** por:

1. Para objetos:
$$\begin{aligned} \mathbf{U}\colon \boldsymbol{Ob}(\mathbf{CPS}) &\longrightarrow \boldsymbol{Ob}(\mathbf{Meas}) \\ (X, \mathcal{F}_X, \mathbb{P}_X) &\longmapsto (X, \mathcal{F}_X). \end{aligned}$$

2. Sean $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ y $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ espacios de probabilidad, para flechas:

$$\begin{aligned} \mathbf{U}\colon \mathbf{CPS}(\overline{X}, \overline{Y}) &\longrightarrow \mathbf{Meas}((X, \mathcal{F}_X), (Y, \mathcal{F}_Y)) \\ [f\colon \overline{X} \to \overline{Y}] &\longmapsto [f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)]. \end{aligned}$$

En este caso, el funtor $\mathcal{U}$ envía un espacio de probabilidad $\overline{X}$ en un espacio medible $(X, \mathcal{F}_X)$ olvidando su medida de probabilidad.

En general, no es posible definir formalmente un funtor de olvido; sin embargo, la idea de este funtor es que *olvida* alguna estructura con respecto a los objetos y flechas de la categoría de dominio.

- Entre las categorías **CMS** y **CPS** se define el funtor de normalización canónica $\mathcal{N}\colon \textbf{CMS} \to \textbf{CPS}$ por:

  1. Para objetos:

$$\begin{aligned} \mathcal{N}\colon \textbf{\textit{Ob}}(\textbf{CMS}) &\longrightarrow \textbf{\textit{Ob}}(\textbf{CPS}) \\ (X, \mathcal{F}_X, \mu) &\longmapsto \left(X, \mathcal{F}_X, \tfrac{1}{\mu(X)}\mu\right). \end{aligned}$$

  2. Sean $\overline{X} = (X, \mathcal{F}_X, \mu)$ y $\overline{Y} = (Y, \mathcal{F}_Y, \nu)$ espacios medidos finitos, para flechas:

$$\begin{aligned} \mathcal{N}\colon \textbf{CMS}(\overline{X}, \overline{Y}) &\longrightarrow \textbf{CPS}\left(\left(X, \mathcal{F}_X, \tfrac{1}{\mu(X)}\mu\right), \left(Y, \mathcal{F}_Y, \tfrac{1}{\nu(Y)}\nu\right)\right) \\ [f\colon \overline{X} \to \overline{Y}] &\longmapsto \left[\mathcal{N}(f)\colon \left(X, \mathcal{F}_X, \tfrac{1}{\mu(X)}\mu\right) \to \left(Y, \mathcal{F}_Y, \tfrac{1}{\nu(Y)}\nu\right)\right]. \end{aligned}$$

  Vemos que el funtor $\mathcal{N}$ envía un espacio medido finito en un espacio de probabilidad. Además, notemos que $\mathcal{N}(f)$ es una función acotada. En efecto, dado que f es una función acotada, sabemos existe $M > 0$ tal que $\mu(f^{-1}(F_Y)) \leq M\nu(F_Y)$ para todo $F_Y \in \mathcal{F}_Y$. Así, tomando $M_1 = \frac{\nu(Y)}{\mu(X)}M > 0$, se tiene que

$$\frac{1}{\mu(X)}\mu(f^{-1}(F_Y)) \leq M_1 \frac{1}{\nu(Y)}\nu(F_Y),$$

  para todo $F_Y \in \mathcal{F}_Y$.

Presentaremos el funtor de Giry definido en [3], para ello, recordemos la categoría **Meas** de espacios medibles y funciones medibles.

DEFINICIÓN 2.17 (Funtor de Giry)**.** *Sean $(X, \mathcal{F}_X)$ y $(Y, \mathcal{F}_Y)$ dos espacios medibles. El funtor de Giry $\mathcal{P}\colon$ **Meas** $\to$ **Meas** está definido por:*

1. *Para objetos:*

$$\begin{aligned} \mathcal{P}\colon \textbf{\textit{Ob}}(\textbf{\textit{Meas}}) &\longrightarrow \textbf{\textit{Ob}}(\textbf{\textit{Meas}}) \\ (X, \mathcal{F}_X) &\longmapsto \mathcal{P}(X, \mathcal{F}_X) := (\textbf{\textit{P}}(X), \textbf{\textit{P}}(\mathcal{F}_X)) \end{aligned}$$

   donde $\textbf{\textit{P}}(X)$ es el conjunto de medidas de probabilidad definidas so-

bre X, es decir,

$$\boldsymbol{P}(X) := \{\mathbb{P}_X \colon (X, \mathcal{F}_X) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) : \mathbb{P}_X \text{ es una medida de probabilidad}\}.$$

Por otro lado, para definir la $\sigma-$álgebra sobre $\boldsymbol{P}(X)$, sea $F_X \in \mathcal{F}_X$, consideremos la función de evaluación definida por

$$\begin{aligned} i_{F_X} \colon \boldsymbol{P}(X) &\longrightarrow [0,1] \\ \mathbb{P}_X &\longmapsto i_{F_X}(\mathbb{P}_X) = \mathbb{P}_X(F_X). \end{aligned}$$

Gracias a la definición A.2, podemos considerar la $\sigma-$álgebra más pequeña tal que la familia de funciones $\{i_{F_X}\}_{F_X \in \mathcal{F}_X}$ es medible. Así, tenemos que

$$\begin{aligned} \boldsymbol{P}(\mathcal{F}_X) &:= \sigma\left(\{i_{F_X}^{-1}(B) : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\}\right), \\ &= \sigma\left(\{\{\mathbb{P}_X \in \boldsymbol{P}(X) : \mathbb{P}_X(F_X) \in B\} : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\}\right). \end{aligned}$$

Con lo cual, obtenemos el espacio medible $(\boldsymbol{P}(X), \boldsymbol{P}(\mathcal{F}_X))$.

2. *Para flechas:*

$$\begin{aligned} \mathcal{P} \colon \textbf{Meas}((X,\mathcal{F}_X),(Y,\mathcal{F}_Y)) &\longrightarrow \textbf{Meas}((\boldsymbol{P}(X),\boldsymbol{P}(\mathcal{F}_X)),(\boldsymbol{P}(Y),\boldsymbol{P}(\mathcal{F}_Y))) \\ [f \colon (X,\mathcal{F}_X) \to (Y,\mathcal{F}_Y)] &\longmapsto [\mathcal{P}(f) \colon (\boldsymbol{P}(X),\boldsymbol{P}(\mathcal{F}_X)) \to (\boldsymbol{P}(Y),\boldsymbol{P}(\mathcal{F}_Y))], \end{aligned}$$

donde $\boldsymbol{P}(f)$ es una función tal que

$$\begin{aligned} \mathcal{P}(f) \colon (\boldsymbol{P}(X),\boldsymbol{P}(\mathcal{F}_X)) &\longrightarrow (\boldsymbol{P}(Y),\boldsymbol{P}(\mathcal{F}_Y)) \\ \mathbb{P}_X &\longmapsto \mathbb{P}_X \circ f^{-1}, \end{aligned}$$

donde $\mathbb{P} \circ f^{-1}$ es la imagen de la medida $\mathbb{P}_X$ con respecto a la función f.

PROPOSICIÓN 2.2. *$\mathcal{P}$ es un funtor sobre la categoría **Meas**.*

Demostración. La demostración la haremos en dos pasos, primero probaremos que para una flecha $f \colon (X,\mathcal{F}_X) \to (Y,\mathcal{F}_Y)$ en **Meas** la función $\mathcal{P}(f) \colon (\mathbf{P}(X),\mathbf{P}(\mathcal{F}_X)) \to (\mathbf{P}(Y),\mathbf{P}(\mathcal{F}_Y))$ es $\mathbf{P}(\mathcal{F}_X)/\mathbf{P}(\mathcal{F}_Y)-$medible. Luego, probaremos que el functor de Giry preserva las composiciones y las identidades.

1. Sea $A \in \{i_F^{-1}(B) : F_Y \in \mathcal{F}_Y \text{ y } B \in \mathcal{B}(\mathbb{R})\}$, sabemos existen $F_Y \in \mathcal{F}_Y$ y $B \in \mathcal{B}(\mathbb{R})$ tales que
$$A = i_F^{-1}(B).$$
Gracias al lema A.2, para probar que $\mathcal{P}(f)$ es $\mathbf{P}(\mathcal{F}_X)/\mathbf{P}(\mathcal{F}_Y)-$medible, basta con demostrar que $(\mathcal{P}(f))^{-1}(i_F^{-1}(B)) \in \sigma(\{i_{F_X}^{-1}(B) : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\})$. En efecto, tenemos que
$$\begin{aligned}(\mathcal{P}(f))^{-1}(i_{F_Y}^{-1}(B)) &= \{\mathbb{P}_X \in \mathbf{P}(X) : \mathcal{P}(f)(\mathbb{P}_X) \in i_{F_Y}^{-1}(B)\},\\ &= \{\mathbb{P}_X \in \mathbf{P}(X) : \mathbb{P}_X \circ f^{-1} \in i_{F_Y}^{-1}(B)\},\\ &= \left\{\mathbb{P}_X \in \mathbf{P}(X) : i_{F_Y}\left(\mathbb{P}_X \circ f^{-1}\right) \in B\right\},\\ &= \left\{\mathbb{P}_X \in \mathbf{P}(X) : \mathbb{P}_X\left(f^{-1}(F_Y)\right) \in B\right\},\\ &= i_{f^{-1}(F_Y)}^{-1}(B),\end{aligned}$$
de donde, dado que $f^{-1}(F_Y) \in \mathcal{F}_X$, se sigue que
$$(\mathcal{P}(f))^{-1}(i_{F_Y}^{-1}(B)) \in \sigma(\{i_{F_X}^{-1}(B) : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\}),$$
como queríamos.

   Por lo tanto, podemos concluir que $\mathcal{P}(f)\colon (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X)) \to (\mathbf{P}(Y), \mathbf{P}(\mathcal{F}_Y))$ es una función $\mathbf{P}(\mathcal{F}_X)/\mathbf{P}(\mathcal{F}_Y)-$medible

2. Vamos a demostrar que $\mathcal{P}$ preserva la composición y las identidades.

   - **Asociatividad:** Sean $f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ y $g\colon (Y, \mathcal{F}_Y) \to (Z, \mathcal{F}_Z)$ flechas componibles en **Meas**, vamos a demostrar que
$$\mathcal{P}(g \circ f) = \mathcal{P}(g) \circ \mathcal{P}(f).$$
Por un lado, sea $\mathbb{P}_X \in \mathbf{P}(X)$, aplicando la definición de $\mathcal{P}$ para la flecha $g \circ f\colon (X, \mathcal{F}_X) \to (Z, \mathcal{F}_Z)$, tenemos que
$$\mathcal{P}(g \circ f)(\mathbb{P}_X) = \mathbb{P}_X \circ (g \circ f)^{-1}.$$
Ahora, sea $F_Z \in \mathcal{F}_Z$, por definición de imagen de la medida $\mathbb{P}_X$ con respecto a la función $g \circ f$, sabemos que
$$\left(\mathbb{P}_X \circ (g \circ f)^{-1}\right)(F_Z) = \mathbb{P}_X\left((g \circ f)^{-1}(F_Z)\right),$$

con lo cual, se tiene que

$$\mathbb{P}_X\left((g\circ f)^{-1}(F_Z)\right)=\mathbb{P}_X\left(f^{-1}(g^{-1}(F_Z))\right). \tag{2.9}$$

Por otro lado, sabemos que

$$\mathcal{P}(f)(\mathbb{P}_X)=\mathbb{P}_X\circ f^{-1},$$

con lo cual, dado que $\mathbb{P}_X\circ f^{-1}\in\mathbf{P}(Y)$, se sigue que

$$\mathcal{P}(g)\left(\mathbb{P}_X\circ f^{-1}\right)=\left(\mathbb{P}_X\circ f^{-1}\right)\circ g^{-1}.$$

Luego, sea $F_Z\in\mathcal{F}_Z$, por definición de imagen de la medida $\mathbb{P}_X\circ f^{-1}$ con respecto a la función g, se tiene que

$$\left(\left(\mathbb{P}_X\circ f^{-1}\right)\circ g^{-1}\right)(F_Z)=(\mathbb{P}_X\circ f^{-1})(g^{-1}(F_Z)),$$

con esto, nuevamente aplicando la definición imagen de la medida $\mathbb{P}_X$ con respecto a la función f, obtenemos que

$$\left(\left(\mathbb{P}_X\circ f^{-1}\right)\circ g^{-1}\right)(F_Z)=\mathbb{P}_X\left(f^{-1}(g^{-1})(F_Z)\right). \tag{2.10}$$

Combinando (2.9) y (2.10), podemos concluir que

$$\left(\mathbb{P}_X\circ(g\circ f)^{-1}\right)(F_Z)=\left(\left(\mathbb{P}_X\circ f^{-1}\right)\circ g^{-1}\right)(F_Z),$$

para cada $F\in\mathcal{F}_Z$.
Es decir, hemos probado que

$$\mathcal{P}(g\circ f)(\mathbb{P}_X)=(\mathcal{P}(g)\circ\mathcal{P}(f))(\mathbb{P}_X),$$

para cada $\mathbb{P}_X\in\mathbf{P}(X)$, con lo cual podemos concluir que

$$\mathcal{P}(g\circ f)=\mathcal{P}(g)\circ\mathcal{P}(f),$$

como queríamos.

- **Identidades:**

  Finalmente, consideremos la flecha identidad $Id\colon(X,\mathcal{F}_X)\to(X,\mathcal{F}_X)$, vamos a demostrar que

$$\mathcal{P}(Id) = I,$$

donde $I\colon (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X)) \to (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X))$ es tal que $I(\mathbb{P}_X) = \mathbb{P}_X$ para cada $\mathbb{P}_X \in \mathbf{P}(X)$.
Para ello, sea $\mathbb{P}_X \in \mathbf{P}(X)$, se tiene que

$$\mathcal{P}(Id)(\mathbb{P}_X) = \mathbb{P}_X \circ Id^{-1},$$

de donde, dado que $\mathbb{P}_X(Id^{-1}(A)) = \mathbb{P}_X(A)$ para todo $A \in \mathcal{F}_X$, se sigue que

$$\mathcal{P}(Id)(\mathbb{P}_X) = \mathbb{P}_X,$$

con esto, podemos concluir que

$$\mathcal{P}(Id) = I. \qquad \square$$

2.1.3. Transformaciones naturales

Las transformaciones naturales son una noción de "mapa entre funtores" y se aplica cuando los funtores tienen el mismo dominio y codominio. Formalmente, tenemos la siguiente definición:

Definición 2.18 (Transformación natural)**.** *Sean $\mathcal{C}, \mathcal{D}$ categorías y $F\colon \mathcal{C} \to \mathcal{D}$, $G\colon \mathcal{C} \to \mathcal{D}$ funtores. Una* transformación natural $\alpha\colon F \to G$ *es una familia*
$(\alpha_A\colon F(A) \to G(A))_{A \in \mathbf{Ob}(\mathcal{C})}$ de flechas en $\mathcal{D}$ tal que, para cada flecha $f\colon A \to B$ en $\mathcal{C}$, el diagrama

$$\begin{tikzcd}
F(A) \arrow[r, "F(f)"] \arrow[d, "\alpha_A"'] & F(B) \arrow[d, "\alpha_B"] \\
G(A) \arrow[r, "G(f)"'] & G(B)
\end{tikzcd}$$

conmuta, es decir,

$$\alpha_B \circ F(f) = G(f) \circ \alpha_A. \tag{2.11}$$

Las flechas α_A se denominan las componentes de la transformación natural α.

Observación. La definición de transformación natural 2.18 establece que para cada flecha $f \colon A \to B$ en $\mathcal{C}$ es posible construir exactamente una flecha $: F(A) \to G(B)$ en $\mathcal{D}$.

Además, la identidad (2.11) se denomina condición de naturalidad.

Podemos ver que las transformaciones naturales son un tipo de flecha, por lo que esperaríamos poder componerlas. Para ello, sean $\mathcal{C}, \mathcal{D}$ categorías, $F, G, H \colon \mathcal{C} \to \mathcal{D}$ functores y $\alpha \colon F \to G$, $\beta \colon G \to H$ transformaciones naturales. La transformación natural compuesta $\beta \circ \alpha \colon F \to H$ esta definida por

$$(\beta \circ \alpha)_A = \beta_A \circ \alpha_A,$$

para cada $A \in \textbf{\textit{Ob}}(\mathcal{C})$. También podemos considerar la transformación natural identidad $1_F \colon F \to F$, dada por

$$(1_F)_A = 1_{F(A)},$$

para cada $A \in \textbf{\textit{Ob}}(\mathcal{C})$. Con esto, para dos categorías $\mathcal{C}$ y $\mathcal{D}$, existe una categoría cuyos objetos son los funtores de la categoría $\mathcal{C}$ a la categoría $\mathcal{D}$ y cuyas flechas entre funtores son las transformaciones naturales. Denotaremos a esta categoría por $[\mathcal{C}, \mathcal{D}]$ y la llamaremos la categoría funtor de $\mathcal{C}$ a $\mathcal{D}$.

Así, presentamos un ejemplo de transformación natural empleando herramientas de teoría de probabilidades:

Ejemplo 13. Sean $(X, \mathcal{F}_X)$ un espacio medible y la función $\alpha_{(X,\mathcal{F}_X)}$ definida por

$$\begin{aligned} \alpha_{(X,\mathcal{F}_X)} \colon (X, \mathcal{F}_X) &\longrightarrow (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X)) \\ x &\longmapsto \alpha_X(x) = \delta_x, \end{aligned}$$

donde δ_x es la medida de Dirac en el punto $x \in X$.

Además, sean $1_{\mathbf{Meas}} \colon \mathbf{Meas} \to \mathbf{Meas}$ el funtor identidad en **Meas** y el funtor de Giry $\mathcal{P}\colon \mathbf{Meas} \to \mathbf{Meas}$.

Se tiene que $\alpha\colon 1_{\mathbf{Meas}} \to \mathcal{P}$ es una transformación natural. La demostración del hecho de que α es una transformación natural la haremos en dos pasos:

1. Sea $(X, \mathcal{F}_X) \in \boldsymbol{Ob}(\mathbf{Meas})$ un objeto en **Meas**, vamos a demostrar que $\alpha_{(X,\mathcal{F}_X)}\colon (X, \mathcal{F}_X) \to (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X))$ es una función $\mathcal{F}_X/\mathbf{P}(\mathcal{F}_X)-$ medible.

   Para ello, sea $A \in \{i_{F_X}^{-1}(B) : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\}$, sabemos existen $F_X \in \mathcal{F}_X$ y $B \in \mathcal{B}(\mathbb{R})$ tales que
$$A = i_{F_X}^{-1}(B).$$
   Por el lema A.2, para probar la medibilidad de $\alpha_{(X,\mathcal{F}_X)}$ basta con demostrar que $\alpha_{(X,\mathcal{F}_X)}^{-1}(A) \in \mathcal{F}_X$. En efecto, se tiene que
$$\begin{aligned}
\alpha_{(X,\mathcal{F}_X)}^{-1}(A) &= \alpha_{(X,\mathcal{F}_X)}^{-1}\left(i_{F_X}^{-1}(B)\right), \\
&= \{x \in X : \alpha(x) \in i_{F_X}^{-1}(B)\}, \\
&= \{x \in X : i_{F_X}(\alpha_x) \in B\}, \\
&= \{x \in X : \alpha_x(F_X) \in B\}.
\end{aligned} \tag{2.12}$$
   Por otro lado, sabemos que
$$\alpha_x(F_X) = \begin{cases} 1 & \text{si} \quad x \in F_X, \\ 0 & \text{si} \quad x \in F_X^C. \end{cases}$$
   Así, junto con (2.12), consideremos los siguientes casos:

   a) Si $\{1\} \in B$, entonces $\alpha_{(X,\mathcal{F}_X)}^{-1}(A) = F_X$.

   b) Si $\{0\} \in B$, entonces $\alpha_{(X,\mathcal{F}_X)}^{-1}(A) = F_X^C$.

   c) Si $\{0, 1\} \in B$, entonces $\alpha_{(X,\mathcal{F}_X)}^{-1}(A) = X$.

   d) Si $\{0, 1\} \notin B$, entonces $\alpha_{(X,\mathcal{F}_X)}^{-1}(A) = \emptyset$.

   Podemos ver que en todos los casos se tiene que $\alpha_{(X,\mathcal{F}_X)}^{-1}(A) \in \mathcal{F}_X$. Por lo tanto, podemos concluir que $\alpha_{(X,\mathcal{F}_X)}\colon (X, \mathcal{F}_X) \to (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X))$ es una función $\mathcal{F}_X/\mathbf{P}(\mathcal{F}_X)-$ medible.

2. Ahora, probaremos que se satisface la condición de naturalidad. En efecto, dado que $1_{\mathbf{Meas}}(X, \mathcal{F}_X) = (X, \mathcal{F}_X)$ para todo objeto $(X, \mathcal{F}_X) \in$ ***Ob***(**Meas**) y $1_{\mathbf{Meas}}(f) = f$ toda flecha $f \in \boldsymbol{\mathcal{A}}(\mathbf{Meas})$, se debe probar que el siguiente diagrama

$$\begin{tikzcd}
(X, \mathcal{F}_X) \arrow[r, "f"] \arrow[d, "\alpha_X"'] & (Y, \mathcal{F}_Y) \arrow[d, "\alpha_Y"] \\
(\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X)) \arrow[r, "\mathcal{P}(f)"'] & (\mathbf{P}(Y), \mathbf{P}(\mathcal{F}_Y))
\end{tikzcd}$$

conmuta, para lo cual, vamos a verificar que

$$\alpha_Y \circ f = \mathcal{P}(f) \circ \alpha_X.$$

Sea $x \in X$, por un lado, se tiene que

$$\begin{aligned}(\alpha_Y \circ f)(x) &= \alpha_Y(f(x)) \\ &= \delta_{f(x)},\end{aligned}$$

de donde, para cada $A \in \mathcal{F}_Y$, por definición de la medida de Dirac en el punto $f(x) \in A$, obtenemos que

$$\delta_{f(x)}(A) = \begin{cases} 1 & \text{si} \quad f(x) \in A, \\ 0 & \text{si} \quad f(x) \in A^C. \end{cases} = \begin{cases} 1 & \text{si} \quad x \in f^{-1}(A), \\ 0 & \text{si} \quad x \in f^{-1}(A^C). \end{cases} \tag{2.13}$$

Por otro lado, tenemos que

$$\begin{aligned}\left(\mathcal{P}(f) \circ \alpha_X\right)(x) &= \mathcal{P}(f)(\delta_x), \\ &= \delta_x \circ f^{-1},\end{aligned}$$

de donde, para cada $A \in \mathcal{F}_Y$, se sigue que

$$\left(\delta_x \circ f^{-1}\right)(A) = \delta_x(f^{-1}(A)) = \begin{cases} 1 & \text{si} \quad x \in f^{-1}(A), \\ 0 & \text{si} \quad x \in f^{-1}(A^C). \end{cases} \tag{2.14}$$

Así, junto con (2.13) y (2.14), podemos concluir que

$$\alpha_Y \circ f = \mathcal{P}(f) \circ \alpha_X,$$

como queríamos.

2.1.4. Productos

Definición 2.19. *Sean $\mathcal{C}$ una categoría y $A, B \in$ **Ob**$(\mathcal{C})$ objetos. Un* producto *de A y B consiste de:*

- *Un objeto $P \in$ **Ob**$(\mathcal{C})$;*
- *Un par de flechas $p_1 \colon P \to A$ y $p_2 \colon P \to B \in \boldsymbol{\mathcal{A}}(\mathcal{C})$;*

*tales que, para cada objeto $C \in$ **Ob**$(\mathcal{C})$ y para cada par de flechas $f_1 \colon C \to A$ y $f_2 \colon C \to B$, existe una única flecha $\overline{f} \colon C \to P$ en $\mathcal{C}$ tal que el siguiente diagrama:*

$$\begin{tikzcd}
& C \arrow[dl, "f_1"'] \arrow[d, dashed, "\overline{f}"] \arrow[dr, "f_2"] & \\
A & P \arrow[l, "p_1"] \arrow[r, "p_2"'] & B
\end{tikzcd}$$

conmuta, es decir,

$$p_1 \circ \overline{f} = f_1 \quad \text{y} \quad p_2 \circ \overline{f} = f_2.$$

Además, las flechas p_1 y p_2 se denominan proyecciones.

Observación.

- Los productos no siempre existen. Como ejemplo, consideremos la Categoría **2**:
$$A \xrightarrow{f} B.$$
Es fácil ver que A y B no tienen producto. Sin embargo, si los objetos A y B de una categoría tiene un producto, entonces es único salvo un único isomorfismo.

- La tripla (P, p_1, p_2) representará el producto de los objetos A y B.

Como es de costumbre, primero presentemos un ejemplo clásico de producto.

EJEMPLO 14. Consideremos la categoría **Set** y $A, B \in \boldsymbol{Ob}(\mathbf{Set})$ objetos. Recordemos que el producto cartesiano usual de A y B viene dado por $A \times B = \{(a, b) : a \in A \text{ y } b \in B\}$ y las funciones proyecciones se definen por

$$\begin{aligned} p_1 \colon A \times B &\longrightarrow A \\ (a, b) &\longmapsto a \end{aligned} \quad \text{y} \quad \begin{aligned} p_2 \colon A \times B &\longrightarrow B \\ (a, b) &\longmapsto b. \end{aligned}$$

Se tiene que la tripleta $(A \times B, p_1, p_2)$ es un producto en **Set** de los objetos A y B. En efecto, sean $C \in \boldsymbol{Ob}(\mathcal{C})$ y $f \colon C \to A$, $g \colon C \to B$ flechas en **Set**, definamos la flecha $\overline{f}$ por

$$\begin{aligned} \overline{f} \colon C &\longrightarrow A \times B \\ c &\longmapsto (f(c), g(c)). \end{aligned}$$

Así, es fácil ver que el siguiente diagrama

$$\begin{tikzcd} & C \arrow[dl, "f"'] \arrow[d, dashed, "\overline{f}"] \arrow[dr, "g"] & \\ A & A \times B \arrow[l, "p_1"] \arrow[r, "p_2"'] & B \end{tikzcd}$$

conmuta.

También, es fácil ver la función $\overline{f} \colon C \to A \times B$ es única.

Gracias al ejemplo precedente podemos considerar el siguiente producto en el contexto probabilístico:

Ejemplo 15. Consideremos la categoría **Meas** y $(X,\mathcal{F}_X),(Y,\mathcal{F}_Y)\in \boldsymbol{Ob}(\mathbf{Meas})$. Además, sea $(X\times Y,\mathcal{F}_X\otimes\mathcal{F}_Y)$ el espacio medible producto, donde $\mathcal{F}_X\otimes\mathcal{F}_Y$ es la $\sigma-$álgebra más pequeña tal que las funciones proyecciones p_1 y p_2 son medibles. Por otro lado, sean $(Z,\mathcal{F}_Z)\in \boldsymbol{Ob}(\mathbf{Meas})$ y $f\colon(Z,\mathcal{F}_Z)\to(X,\mathcal{F}_X)$, $g\colon(Z,\mathcal{F}_Z)\to(X,\mathcal{F}_X)$ flechas en **Meas**, definamos la función $\overline{f}$ por

$$\begin{aligned}\overline{f}\colon(Z,\mathcal{F}_Z)&\longrightarrow(X\times Y,\mathcal{F}_X\otimes\mathcal{F}_Y)\\ c\quad&\longmapsto(f(c),g(c)).\end{aligned}$$

Se tiene que el siguiente diagrama

$$\begin{tikzcd}
& (Z,\mathcal{F}_Z) \arrow[dl, "f"'] \arrow[d, dashed, "\overline{f}"] \arrow[dr, "g"] & \\
(X,\mathcal{F}_X) & (X\times Y,\mathcal{F}_X\otimes\mathcal{F}_Y) \arrow[l, "p_1"] \arrow[r, "p_2"'] & (Y,\mathcal{F}_Y)
\end{tikzcd}$$

conmuta. Así, para probar que $((X\times Y,\mathcal{F}_X\otimes\mathcal{F}_Y),p_1,p_2)$ es un producto en **Meas**, basta con verificar que $\overline{f}\colon(Z,\mathcal{F}_Z)\to(X\times Y,\mathcal{F}_X\otimes\mathcal{F}_Y)$ es una función $\mathcal{F}_Z/\mathcal{F}_X\otimes\mathcal{F}_Y-$medible. Para ello, sea $F_X\times F_Y\in\mathcal{F}_X\otimes\mathcal{F}_Y$, vamos a demostrar que $\overline{f}^{-1}(F_X\times F_Y)\in\mathcal{F}_Z$.

Sea $x\in\overline{f}^{-1}(F_X\times F_Y)$, tenemos las siguientes equivalencias

$$\begin{aligned}\overline{f}(x)\in F_X\times F_Y&\iff(f(x),g(x))\in F_X\times F_Y\\ &\iff f(x)\in F_X\ \text{ y }\ g(x)\in F_Y\\ &\iff x\in f^{-1}(F_X)\ \text{ y }\ x\in g^{-1}(F_Y),\end{aligned}$$

con lo cual, podemos concluir que

$$\overline{f}^{-1}(F_X\times F_Y)=f^{-1}(F_X)\cap g^{-1}(F_Y). \tag{2.15}$$

Ahora, dado que $F_X\in\mathcal{F}_X$ y $F_Y\in\mathcal{F}_Y$, se sigue que

$$f^{-1}(F_X)\in\mathcal{F}_Z\quad\text{y}\quad g^{-1}(F_Y)\in\mathcal{F}_Z,$$

de donde, junto con (2.15), podemos colegir que

$$\overline{f}^{-1}(F_X \times F_Y) \in \mathcal{F}_Z,$$

como queríamos.

Por lo tanto, $(X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y, p_1, p_2)$ es un producto en **Meas**.

Ahora, presentamos la definición de independencia de dos flechas en la categoría **mpProb.**

DEFINICIÓN 2.20 (Independencia en **mpProb**)**.** *Sean $f\colon \overline{Z} \to \overline{X}$ y $g\colon \overline{Z} \to \overline{Y}$ dos flechas en* ***mpProb****. Decimos que f y g son* independientes *si existe $q\colon \overline{Z} \to \overline{X} \otimes \overline{Y}$ una flecha en* ***mpProb*** *tal que el siguiente diagrama:*

$$\begin{tikzcd}
& \overline{Z} \arrow[dl, "f"'] \arrow[d, dashed, "q"] \arrow[dr, "g"] & \\
\overline{X} & \overline{X} \otimes \overline{Y} \arrow[l, "p_1"] \arrow[r, "p_2"'] & \overline{Y}
\end{tikzcd}$$

conmuta, es decir:

$$q^- \circ p_1^- = f^- \quad \text{y} \quad q^- \circ p_2^- = g^-,$$

donde $p_1\colon X \times Y \to X$ y $p_2\colon X \times Y \to Y$ son las funciones proyecciones.

La siguiente proposición establece que la noción de independencia definida en la categoría **mpProb** coincide exactamente con la definición de independencia clásica empleada en teoría de probabilidades.

PROPOSICIÓN 2.3. *Dos flechas $f\colon \overline{Z} \to \overline{X}$ y $g\colon \overline{Z} \to \overline{Y}$ en* ***mpProb*** *son independientes, si y solo si, para cada par $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, se tiene que*

$$\mathbb{P}_Z\left(f^{-1}(A) \cap g^{-1}(B)\right) = \mathbb{P}_Z\left(f^{-1}(A)\right) \times \mathbb{P}_Z\left(g^{-1}(B)\right).$$

Demostración. Sean $f\colon \overline{Z} \to \overline{X}$ y $g\colon \overline{Z} \to \overline{Y}$ dos flechas en **mpProb**.

- Para la primera implicación, supongamos que existe $q\colon \overline{Z} \to \overline{X} \otimes \overline{Y}$

una flecha en **mpProb** tal que el siguiente diagrama:

$$\begin{tikzcd}
& \overline{Z} \arrow[dl, "f"'] \arrow[d, dashed, "q"] \arrow[dr, "g"] & \\
\overline{X} & \overline{X} \otimes \overline{Y} \arrow[l, "p_1"] \arrow[r, "p_2"'] & \overline{Y}
\end{tikzcd}$$

conmuta, es decir,

$$p_1 \circ q = f \quad \mathbf{y} \quad p_2 \circ q = g,$$

así, podemos considerar la función

$$\begin{aligned} q : \overline{Z} &\longrightarrow \overline{X} \otimes \overline{Y} \\ \omega &\longmapsto (f(\omega), g(\omega)). \end{aligned}$$

Ahora, sean $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, se tiene que

$$\mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) = \mathbb{P}_Z\left((q^{-1}(p_1^{-1})(A)) \cap (q^{-1}(p_2^{-1}))(B)\right),$$

luego, dado que $p_1^{-1}(A) = A \times Y$ y $p_2^{-1}(B) = X \times B$, se sigue que

$$\begin{aligned} \mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) &= \mathbb{P}_Z\left(q^{-1}(A \times Y) \cap q^{-1}(X \times B)\right), \\ &= \mathbb{P}_Z\left(q^{-1}\left((A \times Y) \cap (X \times B)\right)\right), \end{aligned}$$

con esto, puesto que $(A \times Y) \cap (X \times B) = A \cap B$, obtenemos que

$$\mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) = \mathbb{P}_Z\left(q^{-1}\left(A \times B\right)\right). \tag{2.16}$$

Por otro lado, usando el hecho que $q^{-} : \overline{X} \otimes \overline{Y} \to \overline{Z}$ preserva medidas, sabemos que

$$\mathbb{P}_Z \circ q^{-1} = \mathbb{P}_X \otimes \mathbb{P}_Y = \mathbb{P}_X \mathbb{P}_Y,$$

de donde, junto con (2.16), se sigue que

$$\mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) = \mathbb{P}_X(A) \times \mathbb{P}_Y(B). \tag{2.17}$$

Finalmente, dado que $f^{-} : \overline{X} \to \overline{Z}$ y $g^{-} : \overline{Y} \to \overline{Z}$ son flechas en **mp-**

Prob, sabemos que

$$\mathbb{P}_X = \mathbb{P}_Z \circ f^{-1} \quad \text{y} \quad \mathbb{P}_Y = \mathbb{P}_Z \circ g^{-1},$$

luego, junto con (2.17), podemos concluir que

$$\mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) = \mathbb{P}_Z\left(f^{-1}(A)\right) \times \mathbb{P}_Z\left(g^{-1}(B)\right),$$

como queríamos.

- Para el recíproco, supongamos que $\mathbb{P}_Z\left(f^{-1}(A) \cap g^{-1}(B)\right) = \mathbb{P}_Z\left(f^{-1}(A)\right) \cdot \mathbb{P}_Z\left(g^{-1}(B)\right)$ para cada $A \in \mathcal{F}_X$ y para cada $B \in \mathcal{F}_Y$. Vamos a demostrar que existe $q\colon \overline{Z} \to \overline{X} \otimes \overline{Y}$ una flecha en **mpProb** tal que

$$p_1 \circ q = f \quad \text{y} \quad p_2 \circ q = g.$$

En primer lugar, notemos que $f\colon (Z, \mathcal{F}_Z) \to (X, \mathcal{F}_X)$ y $g\colon (Z, \mathcal{F}_Z) \to (Y, \mathcal{F}_Y)$ son flechas en **Meas**, de donde, la tripla $((X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y), p_1, p_2)$ es un producto en **Meas**. Así, sabemos existe

$$\begin{aligned} h \colon (Z, \mathcal{F}_Z) &\longrightarrow (X \times Y, \mathcal{F}_X \times \mathcal{F}_Y) \\ \omega \quad &\longmapsto h(\omega) = (f(\omega), g(\omega)), \end{aligned}$$

la cual es una función $\mathcal{F}_Z/\mathcal{F}_X \times \mathcal{F}_Y$ medible tal que el siguiente diagrama:

$$\begin{tikzcd} & (Z, \mathcal{F}_Z) \arrow[dl, "f"'] \arrow[d, dashed, "h"] \arrow[dr, "g"] & \\ (X, \mathcal{F}_X) & (X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y) \arrow[l, "p_1"] \arrow[r, "p_2"'] & (Y, \mathcal{F}_Y) \end{tikzcd}$$

conmuta, es decir,

$$p_1 \circ h = f \quad \text{y} \quad p_2 \circ h = g.$$

Así, probemos que $h\colon \overline{Z} \to \overline{X} \otimes \overline{Y}$ es una flecha en **mpProb**, para lo cual, basta demostrar que $\mathbb{P}_X \otimes \mathbb{P}_Y = \mathbb{P}_Z \circ h^{-1}$. En efecto, sean $A \in \mathcal{F}_X$

y $B \in \mathcal{F}_Y$, tenemos que

$$\mathbb{P}_Z\left(h^{-1}(A \times B)\right) = \mathbb{P}_Z\left(f^{-1}(A) \cap g^{-1}(B)\right),$$

de donde, junto con la hipótesis obtenemos que

$$\mathbb{P}_Z\left(h^{-1}(A \times B)\right) = \mathbb{P}_Z(f^{-1}(A))\mathbb{P}_Z(g^{-1}(B)),$$

luego, dado que $\mathbb{P}_X = \mathbb{P}_Z \circ f^{-1}$ y $\mathbb{P}_Y = \mathbb{P}_Z \circ g^{-1}$, se sigue que

$$\mathbb{P}_Z\left(h^{-1}(A \times B)\right) = \mathbb{P}_X(A)\mathbb{P}_Y(B) = \mathbb{P}_X \otimes \mathbb{P}_Y(A \times B),$$

para todo $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, con lo cual podemos concluir que

$$\mathbb{P}_X \otimes \mathbb{P}_Y = \mathbb{P}_Z \circ h^{-1},$$

es decir, hemos probado que $h\colon \overline{Z} \to \overline{X} \otimes \overline{Y}$ es una flecha en **mpProb**.

$\square$

2.1.5. Mónadas

Consideremos $(X, \mathcal{F})$ un espacio medible y denotemos por $\mathbf{P}(X)$ al espacio de medidas de probabilidad sobre X. Las medidas de probabilidad en un espacio X pueden considerarse como leyes de variables aleatorias. Un tópico de interés en Probabilidad y Estadística es conocer variables aleatorias cuya ley también es aleatoria, es decir, que su ley esté en el espacio $\mathbf{P}(\mathbf{P}(X))$ (el espacio de medidas de probabilidad sobre $\mathbf{P}(X)$).

En general, el espacio de las medidas de probabilidad $\mathbf{P}(X)$ es más grande que el espacio X. A su vez, el espacio $\mathbf{P}(\mathbf{P}(X))$ de las medidas de probabilidad sobre el espacio de las distribuciones de probabilidad es aún más grande, esto hace que en la práctica sea difícil trabajar con el mismo. Por ejemplo, $\mathbf{P}(\mathbf{P}(X))$ puede ser de dimensión infinita a pesar de que X sea finito. Este problema se puede tratar por medio de la estadística paramétrica, otra perspectiva consiste en trabajar en un entorno donde el espacio de medidas de probabilidad $\mathbf{P}(X)$ hereda varias de las propiedades del espacio X, por lo que estos espacios se pueden estudiar de la misma manera, este enfoque se aborda en [14].

En [9] se prueba que estos enfoques son equivalentes a emplear mónadas de probabilidad, debido a que una mónada nos permite tratar construcciones recursivas como la de $X, \mathbf{P}(X)$ y $\mathbf{P}(\mathbf{P}(X))$. El concepto más central en probabilidad categórica es el de mónada de probabilidad, la primera mónada de probabilidad es la mónada de Giry presentada en el Ejemplo 2.22 y las primeras ideas sobre su estructura se encuentran en Lawvere [6].

Previo a definir el concepto de mónada, fijemos cierta notación. Sean $\mathcal{C}, \mathcal{D}, \mathcal{E}$ categorías, $F, G \colon \mathcal{C} \to \mathcal{D}, H \colon \mathcal{D} \to \mathcal{E}$ functores. Si $\eta \colon F \to G$ es una transformación natural, entonces se puede definir la transformación natural $H\eta \colon H \circ F \to H \circ G$ por

$$(H\eta)_A = H(\eta_A),$$

para cada $A \in \textbf{\textit{Ob}}(\mathcal{A})$.

Por otro lado, si $K \colon \mathcal{E} \to \mathcal{D}$ es un functor se puede definir la transformación natural $\eta K \colon F \circ K \to G \circ K$ por

$$(\eta K)_A = \eta_{K(A)},$$

para cada $A \in \textbf{\textit{Ob}}(\mathcal{E})$.

Definición 2.21 (Mónada)**.** *Sea $\mathcal{C}$ una categoría. Una* mónada *en $\mathcal{C}$ consiste de:*

- *Un funtor $T \colon \mathcal{C} \to \mathcal{C}$;*
- *Una transformación natural $\eta \colon 1_{\mathcal{C}} \to T$, llamada unidad;*
- *Una transformación natural $\mu \colon T \circ T \to T$, llamada composición o multiplicación;*

tales que los siguientes diagramas conmutan:

Estos diagramas se denominan unidad izquierda, unidad derecha y asociatividad, respectivamente.

Observación. Podemos reescribir el diagrama en términos de las componentes de las transformaciones naturales. Para ello, sea $A \in \textbf{\textit{Ob}}(\mathcal{C})$,

$$
\begin{tikzcd}
T \arrow[r, "\eta T"] \arrow[dr, "\mathbf{1}_T"'] & T \circ T \arrow[d, "\mu"] \\
& T
\end{tikzcd}
\quad
\begin{tikzcd}
T \arrow[r, "T\eta"] \arrow[dr, "\mathbf{1}_T"'] & T \circ T \arrow[d, "\mu"] \\
& T
\end{tikzcd}
\quad
\begin{tikzcd}
T \circ T \circ T \arrow[r, "T\mu"] \arrow[d, "\mu T"] & T \circ T \arrow[d, "\mu"] \\
T \circ T \arrow[r, "\mu"'] & T
\end{tikzcd}
$$

tenemos que la unidad $\eta_A \colon A \to T(A)$ y la composición $\mu_A \colon T(T(A)) \to T(A)$ son flechas en $\mathcal{C}$ tales que los siguientes diagramas conmutan:

$$
\begin{tikzcd}
T(A) \arrow[r, "\eta_{T(A)}"] \arrow[dr, "\mathbf{1}_{T(A)}"'] & T \circ T(A) \arrow[d, "\mu_A"] \\
& T(A)
\end{tikzcd}
\quad
\begin{tikzcd}
T(A) \arrow[r, "T(\eta_A)"] \arrow[dr, "\mathbf{1}_{T(A)}"'] & T \circ T(A) \arrow[d, "\mu_A"] \\
& T(A)
\end{tikzcd}
\quad
\begin{tikzcd}
T \circ T \circ T(A) \arrow[r, "T(\mu_A)"] \arrow[d, "\mu_{T(A)}"] & T \circ T(A) \arrow[d, "\mu_A"] \\
T \circ T(A) \arrow[r, "\mu_A"'] & T(A)
\end{tikzcd}
$$

Expondremos un ejemplo de mónada en la categoría **Meas**.

DEFINICIÓN 2.22 (Mónada de Giry)**.** *La mónada de Giry está dada por:*

- *El funtor de Giry $\mathcal{P}\colon$ **Meas** $\to$ **Meas** definido en 2.17;*
- ***Unidad**: La transformación natural $\alpha\colon 1_{\textbf{Meas}} \to \mathcal{P}$ presentada en el ejemplo 13;*
- ***Composición:** La transformación natural $E\colon \mathcal{P} \circ \mathcal{P} \to \mathcal{P}$ definida por:*

  Sea $(X, \mathcal{F}_X)$ un espacio medible, para $\pi' \in \mathbf{P}(\mathbf{P}(X))$, se define una medida sobre $(X, \mathcal{F}_X)$ por

  $$E_{(X,\mathcal{F}_X)}(\pi')(F_X) = \int_{\mathbf{P}(X)} i_{F_X}(\mathbb{P}_X)\,\pi'(d\mathbb{P}_X) = \int_{\mathbf{P}(X)} \mathbb{P}_X(F_X)\,\pi'(d\mathbb{P}_X),$$

  para cada $F \in \mathcal{F}_X$.

OBSERVACIÓN. Notemos que $\mathbb{P}_X$ de la Definición 2.22 corresponde a una medida de probabilidad definida sobre $(X, \mathcal{F}_X)$. Además, la integral está bien definida porque las funciones de evaluación son $\mathbf{P}(\mathcal{F}_X)/\mathcal{B}([0,1])-$medibles.

Observación. Sean $(X, \mathcal{F}_X)$ y $(Y, \mathcal{F}_Y)$ espacios medibles. Probemos que

$$E \colon \mathcal{P} \circ \mathcal{P} \to \mathcal{P},$$

es una transformación natural, para lo cual, debemos demostrar que

$$E_{(X,\mathcal{F}_X)} \colon (\mathcal{P} \circ \mathcal{P})(X, \mathcal{F}_X) \to \mathcal{P}(X, \mathcal{F}_X)$$

es una función medible y que, para cualquier flecha $f \colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ en **Meas**, el siguiente diagrama:

$$\begin{tikzcd}
\mathcal{P}(\mathcal{P}((X,\mathcal{F}_X))) \arrow[r, "\mathcal{P}(\mathcal{P}(f))"] \arrow[d, "E_{(X,\mathcal{F}_X)}"'] & \mathcal{P}(\mathcal{P}((Y,\mathcal{F}_Y))) \arrow[d, "E_{(Y,\mathcal{F}_Y)}"] \\
\mathcal{P}((X,\mathcal{F}_X)) \arrow[r, "\mathcal{P}(f)"'] & \mathcal{P}((Y,\mathcal{F}_Y))
\end{tikzcd}$$

conmuta, lo cual es equivalente a demostrar que

$$E_{(Y,\mathcal{F}_Y)} \circ \mathcal{P}(\mathcal{P}(f)) = \mathcal{P}(f) \circ E_{(X,\mathcal{F}_X)}. \tag{2.18}$$

El hecho de que $E_{(X,\mathcal{F}_X)} \colon (\mathbf{P}(\mathbf{P}(X)), \mathbf{P}(\mathbf{P}(\mathcal{F}_X))) \to (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X))$ es $\mathbf{P}(\mathbf{P}(X))/\mathbf{P}(X)-$medible se prueba en [3] y es análoga a la demostración de la medibilidad de la transformación natural $\alpha \colon 1_{\mathbf{Meas}} \to \mathcal{P}$ presentada en el Ejemplo 13. Notemos que $\mathbf{P}(\mathcal{F}_X)$ y $\mathbf{P}(\mathbf{P}(\mathcal{F}_X))$ corresponden a las $\sigma-$álgebras generadas por las funciones inclusión definidas en 2.17.

Probemos que se satisface la condición de naturalidad (2.18). En efecto, sean $\pi' \in \mathbf{P}(\mathbf{P}(X))$ y $F_Y \in \mathcal{F}_Y$, por un lado, gracias a la fórmula de cambio de variable A.4, tenemos que

$$\begin{aligned}
\left(E_{(Y,F_Y)} \circ \mathcal{P}(\mathcal{P}(f))(\pi')\right)(F_Y) &= \int_{\mathbf{P}(Y)} i_{F_Y} \; d((\mathcal{P}(\mathcal{P})(f))(\pi')) \\
&= \int_{\mathbf{P}(Y)} i_{F_Y} \; d(\pi' \circ (\mathcal{P}(f))^{-1})
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathcal{P}(f)^{-1}(\mathbf{P}(Y))} i_{F_Y} \circ \mathcal{P}(f)\; d\pi' \\
&= \int_{\mathbf{P}(X)} i_{F_Y} \circ \mathcal{P}(f)\; d\pi', \\
&= \int_{\mathbf{P}(X)} i_{F_Y} \circ \mathcal{P}(f)(\mathbb{P}_X)\, \pi'(d\mathbb{P}_X), \\
&= \int_{\mathbf{P}(X)} i_{F_Y} \left(\mathbb{P}_X \circ f^{-1}\right)\; \pi'(d\mathbb{P}_X), \\
&= \int_{\mathbf{P}(X)} \mathbb{P}_X \left(f^{-1}(F_Y)\right)\; \pi'(d\mathbb{P}_X). \qquad (2.19)
\end{aligned}
$$

Por otro lado, se tiene que

$$
\begin{aligned}
(\mathcal{P}(f) \circ E_{(X,\mathcal{F}_X)}(\pi'))(F_Y) &= E_{(X,\mathcal{F}_X)}(\pi')(f^{-1}(F_Y)) \\
&= \int_{\mathbf{P}(X)} i_{f^{-1}(F_Y)}\, d\pi' \\
&= \int_{\mathbf{P}(X)} i_{f^{-1}(F_Y)}(\mathbb{P}_X)\, \pi'(d\mathbb{P}_X) \\
&= \int_{\mathbf{P}(X)} \mathbb{P}_X \left(f^{-1}(F_Y)\right)\; \pi'(d\mathbb{P}_X). \qquad (2.20)
\end{aligned}
$$

Junto con (2.19) y (2.20), podemos concluir que

$$
E_{(Y,\mathcal{F}_Y)} \circ \mathcal{P}(\mathcal{P}(f)) = \mathcal{P}(f) \circ E_{(X,\mathcal{F}_X)},
$$

como queríamos.

Consideremos el siguiente resultado presentado en Giry [theorem 1, [3]], el cual nos permite demostrar que la mónada de Giry está bien definida.

Lema 2.4. *Sean $(X, \mathcal{F}_X)$, $(Y, \mathcal{F}_Y)$ espacios medibles, $\mathbb{P}_X \in \mathbf{P}(X)$, $\pi' \in \mathbf{P}(\mathbf{P}(X))$ y $\theta\colon (X, \mathcal{F}_X) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ una función integrable. Se tienen los siguientes resultados:*

1. $\int_X \theta\, d\alpha_{(X,\mathcal{F}_X)}(x) = \theta(x)$ *para* $x \in X$.
2. *La función* $\xi_\theta\colon (\mathbf{P}(X), \mathbf{P}(\mathcal{F}_X)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *definida por* $\xi_\theta(\mathbb{P}_X) = \displaystyle\int_X \theta\, d\mathbb{P}_X$ *es* $\mathbf{P}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})-$ *medible.*

3. $\displaystyle\int_X \theta(x)\, d(E_{(X,\mathcal{F}_X)}(\pi')) = \int_{\mathbf{P}(X)} \xi_\theta(\mathbb{P}_X)\, \pi'(d\mathbb{P}_X)$.

Demostración. 1. Sea $x \in X$, tenemos que

$$\int_X \theta\, d\alpha_{(X,\mathcal{F}_X)}(x) = \int_X \theta\, d\delta_x,$$

de donde, dado que δ_x es la medida de Dirac en el punto $x \in X$, se sigue que

$$\int_X \theta\, d\alpha_{(X,\mathcal{F}_X)}(x) = \theta(x).$$

2. En primer lugar, probaremos las proposiciones 2 y 3 para funciones indicatrices, sea $F_X \in \mathcal{F}_X$.

   - Consideremos $\theta = \mathbb{1}_{F_Z}$, por un lado, se tiene que

$$\begin{aligned}\xi_\theta(\mathbb{P}_X) &= \int_X \theta\, d\mathbb{P}_X,\\ &= \int_X \mathbb{1}_{F_X}\, d\mathbb{P}_X,\\ &= \int_{F_X} d\mathbb{P}_X,\\ &= \mathbb{P}_X(F_X).\end{aligned} \tag{2.21}$$

     Por otro lado, sabemos que $\mathcal{B}(\mathbb{R}) = \sigma(\tau)$ con $\tau \subset \mathbb{R}$ la colección de todos los conjuntos abiertos en $\mathbb{R}$.

     Sea $Y \in \tau$, vamos a demostrar que ξ_θ es una función $\mathbf{P}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})-$medible, lo cual es equivalente a demostrar que $\xi_\theta^{-1}(Y) \in \sigma(\{i_{F_x}^{-1}(B) : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\})$. Gracias a (2.21), notemos que

$$\begin{aligned}\xi_\theta^{-1}(Y) &= \{\mathbb{P}_X \in \mathbf{P}(X) : \xi_\theta(\mathbb{P}_X) \in Y\},\\ &= \{\mathbb{P}_X \in \mathbf{P}(X) : \mathbb{P}_X(F_X) \in Y\},\\ &= i_{F_X}^{-1}(Y),\end{aligned}$$

     con lo cual dado que $F_X \in \mathcal{F}_X$ y $Y \in \mathcal{B}(\mathbb{R})$, se sigue que

$$\xi_\theta^{-1}(Y) \in \sigma(\{i_{F_x}^{-1}(B) : F_X \in \mathcal{F}_X \text{ y } B \in \mathcal{B}(\mathbb{R})\}),$$

     es decir, hemos probado que ξ_θ es una función $\mathbf{P}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})-$medible.

- De igual forma, sea $\theta = \mathbb{1}_{F_X}$, gracias al resultado precedente, tenemos que $\xi_\theta(\mathbb{P}_X) = \mathbb{P}_X(F_X)$ para toda $\mathbb{P}_X \in \mathbf{P}(X)$, con lo cual, se sigue que

$$\int_{\mathbf{P}(X)} \xi_\theta(\mathbb{P}_X)\, \pi'(d\mathbb{P}_X) = \int_{\mathbf{P}(X)} \mathbb{P}_X(F_X)\, \pi'(d\mathbb{P}_X). \tag{2.22}$$

  Ahora, tenemos que

$$\begin{aligned}
\int_X \theta\, dE_{(X,\mathcal{F}_X)}(\pi') &= \int_X \mathbb{1}_{F_X}\, dE_{(X,\mathcal{F}_X)}(\pi'), \\
&= \int_{F_X} dE_{(X,\mathcal{F}_X)}(\pi'), \\
&= E_{(X,\mathcal{F}_X)}(\pi')(F_X), \\
&= \int_{\mathbf{P}(X)} \mathbb{P}_X(F_X)\ \pi'(d\mathbb{P}_X),
\end{aligned}$$

  de donde, junto con (2.22), podemos concluir que

$$\int_X \theta\, dE_{(X,\mathcal{F}_X)}(\pi') = \int_{\mathbf{P}(X)} \xi_\theta\, \pi'(\mathbb{P}_X).$$

Por la linealidad de la integral, los resultados de *2.* y *3.* se cumplen cuando θ es una función simple. El caso general se sigue por el teorema de la convergencia monótona 1.3 y el hecho de que θ puede ser visto como el límite de una sucesión creciente de funciones simples.

$\square$

Observación. El lema precedente nos permite probar que los siguientes diagramas:

$$\begin{tikzcd}
\mathcal{P} \arrow[r, "\alpha \circ \mathcal{P}"] \arrow[dr, "\mathbf{1}_{\mathcal{P}}"'] & \mathcal{P} \circ \mathcal{P} \arrow[d, "E"] \\
& \mathcal{P}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathcal{P} \arrow[r, "\mathcal{P} \circ \alpha"] \arrow[dr, "\mathbf{1}_{\mathcal{P}}"'] & \mathcal{P} \circ \mathcal{P} \arrow[d, "E"] \\
& \mathcal{P}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathcal{P} \circ \mathcal{P} \circ \mathcal{P} \arrow[r, "\mathcal{P} \circ E"] \arrow[d, "E \circ \mathcal{P}"] & \mathcal{P} \circ \mathcal{P} \arrow[d, "E"] \\
\mathcal{P} \circ \mathcal{P} \arrow[r, "E"'] & \mathcal{P}
\end{tikzcd}$$

conmutan.

Las transformaciones naturales $\eta \circ \mathcal{P}, \mathcal{P} \circ \eta$ y $E \circ \mathcal{P}$, se definen por

$$(\alpha \circ \mathcal{P})_{(X,\mathcal{F}_X)} = \alpha_{\mathcal{P}((X,\mathcal{F}_X))}, (\mathcal{P} \circ \alpha)_{(X,\mathcal{F}_X)} = \mathcal{P}(\alpha_{(X,\mathcal{F}_X)}) \text{ y } (E \circ \mathcal{P})_{(X,\mathcal{F}_X)} = E_{\mathcal{P}((X,\mathcal{F}_X))},$$

para cada $(X, \mathcal{F}_X) \in \boldsymbol{Ob}(\mathbf{Meas})$, respectivamente.

Probemos que los diagramas de unidad izquierda, unidad derecha y de asociatividad conmutan. Sea $(X, \mathcal{F}_X)$ un espacio medible.

1. Gracias a la proposición 1 del lema 2.4, es fácil ver que

$$E_{(X,\mathcal{F}_X)} \circ \delta_{\mathcal{P}((X,\mathcal{F}_X))} = \mathbb{1}_{(X,\mathcal{F}_X)} \quad \text{y} \quad E_{(X,\mathcal{F}_X)} \circ \mathcal{P}(\delta_{(X,\mathcal{F}_X)}) = \mathbb{1}_{(X,\mathcal{F}_X)}.$$

2. Ahora vamos probar que el diagrama de asociatividad conmuta, para ello, sean $\pi'' \in \mathbf{P}(\mathbf{P}(\mathbf{P}(X)))$ y $F_X \in \mathcal{F}_X$, debemos demostrar que

$$E_{(X,\mathcal{F}_X)} \circ E_{\mathcal{P}((X,\mathcal{F}_X))}(\pi'')(F_X) = E_{(X,\mathcal{F}_X)} \circ \alpha_{(\mathcal{P}((X,\mathcal{F}_X)))}(\pi'')(F_X).$$

Por un lado, junto con la fórmula de cambio de variable A.4, obtenemos que

$$\begin{aligned}
(E_{(X,\mathcal{F}_X)} \circ \mathcal{P}(E_{(X,\mathcal{F}_X)})(\pi''))(F_X) &= \int_{\mathbf{P}(X)} i_{F_X}\, d\left(\mathcal{P}(E_{(X,F_X)})(\pi'')\right) \\
&= \int_{\mathbf{P}(X)} i_{F_X}\, d(\pi'' \circ E^{-1}_{(X,\mathcal{F}_X)}), \\
&= \int_{E^{-1}_{(X,\mathcal{F}_X)}(\mathbf{P}(X))} i_{F_X} \circ E_{(X,\mathcal{F}_X)}\, d\pi'', \\
&= \int_{\mathbf{P}(\mathbf{P}(X))} i_{F_X} \circ E_{(X,\mathcal{F}_X)}(\pi')\, \pi''(d\pi'), \\
&= \int_{\mathbf{P}(\mathbf{P}(X))} (E_{(X,\mathcal{F}_X)}(\pi'))(F_X)\, \pi''(d\pi'). \qquad (2.23)
\end{aligned}$$

Ahora, por la definición de $E_{(X,\mathcal{F}_X)}$, tenemos que

$$(E_{(X,\mathcal{F}_X)}(\pi'))(F_X) = \int_{\mathbf{P}(X)} i_{F_X}\, d\pi' = \xi_{i_{F_x}}(\pi'),$$

de donde, junto con (2.23) y y por la proposición 2 del lema 2.4, obtenemos que

$$(E_{(X,\mathcal{F}_X)} \circ \mathcal{P}(E_{(X,\mathcal{F}_X)})(\pi''))(F_X) = \int_{\mathbf{P}(\mathbf{P}(X))} \xi_{i_{F_X}}(\pi')\, \pi''(d\pi'),$$

con lo cual, gracias a la proposición 3 del lema 2.4, se sigue que

$$\begin{aligned}(E_{(X,\mathcal{F}_X)} \circ \mathcal{P}(E_{(X,\mathcal{F}_X)})(\pi''))(F_X) &= \int_{\mathbf{P}(X)} i_{F_X}\, d(E_{(X,\mathcal{F}_X)}(\pi')), \\ &= \int_{\mathbf{P}(X)} i_{F_X}(\mathbb{P}_X)\, (E_{(X,\mathcal{F}_X)}(\pi'))(d\mathbb{P}_X). \quad (2.24)\end{aligned}$$

Por otro lado, por definición de E, tenemos que

$$(E_{(X,\mathcal{F}_X)} \circ E_{\mathcal{P}((X,\mathcal{F}_X))}(\pi''))(F_X) = \int_{\mathbf{P}(X)} i_{F_X}(\mathbb{P}_X)\, (E_{(X,\mathcal{F}_X)}(\pi''))d(\mathbb{P}_X). \quad (2.25)$$

Así, junto con (2.24) y (2.25), podemos concluir que

$$E_{(X,\mathcal{F}_X)} \circ E_{\mathcal{P}(X,\mathcal{F}_X)}(\pi'')(F_X) = E_{(X,\mathcal{F}_X)} \circ \mathcal{P}(E_{(X,\mathcal{F}_X)})(\pi'')(F_X),$$

como queríamos.

2.2. La categoría Prob

Existen pocas categorías con objetos como espacios de probabilidad debido a la dificultad de encontrar una condición apropiada para definir flechas entre ellos. Por ejemplo, como pudimos notar los objetos y flechas de la categoría χ son limitados, por esta razón, no se puede emplear esta categoría como base de una teoría de probabilidades.

Con esto en mente, en este capítulo se describe la categoría **Prob** presentada en [13], la cual tiene por objetos a los espacios de probabilidad y cuyas flechas entre objetos son funciones medibles que preservan conjuntos de medida nula. Una vez introducida esta categoría se presenta la generalización de algunas herramientas clásicas de la teoría de la probabilidad, como lo son, por ejemplo, la esperanza condicional, la independencia y la medibilidad; y, su interacción.

Como lo veremos posteriormente las flechas en la categoría **Prob** son adecuadas, por ello, en [12] sobre esta categoría se generalizan los conceptos de filtraciones, procesos adaptados y martingalas.

2.2.1. Definición de la categoría Prob

En primer lugar, vamos a presentar definiciones y resultados que nos permitirán describir la categoría **Prob**. Para ello, consideraremos a $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ como un espacio de probabilidad, es decir, X es el espacio muestral, $\mathcal{F}_X$ una $\sigma-$álgebra de conjuntos de X y $\mathbb{P}_X$ es una medida de probabilidad.

Para definir las flechas en **Prob** consideremos la siguiente definición:

DEFINICIÓN 2.23 (Función nula-conservadora). *Sean $\overline{X}$ y $\overline{Y}$ espacios de probabilidad, una función medible $f\colon \overline{X} \to \overline{Y}$ se dice* nula-conservadora *si $f^{-1}(A) \in \mathbb{P}_Y^{-1}(0) \subset \mathcal{F}_Y$ para cada $A \in \mathbb{P}_X^{-1}(0) \subset \mathcal{F}_X$.*

OBSERVACIÓN. La definición 2.23 establece que una función medible $f\colon \overline{X} \to \overline{Y}$ es nula-conservadora si la pre-imagen bajo f de un conjunto de medida nula es un conjunto de medida nula. En efecto, sea $A \in \mathbb{P}_X^{-1}(0)$, por definición de imagen inversa, tenemos que

$$\mathbb{P}_X(A) = 0.$$

Análogamente, para $f^{-1}(A) \in \mathbb{P}_Y^{-1}(0)$, se tiene que

$$\mathbb{P}_Y\left(f^{-1}(A)\right) = 0.$$

Consideremos la siguiente caracterización sobre la definición de función nula-conservadora:

PROPOSICIÓN 2.5. *Sean $\overline{X}$ y $\overline{Y}$ espacios de probabilidad, una función medible $f\colon \overline{Y} \to \overline{X}$ es nula-conservadora si y solo si $\mathbb{P}_Y \circ f^{-1} \ll \mathbb{P}_X$.*

Demostración.

- Para la primera implicación, supongamos que $f\colon \overline{X} \to \overline{Y}$ es nula-conservadora, vamos a demostrar que $\mathbb{P}_Y \circ f^{-1}$ es absolutamente continua con respecto a $\mathbb{P}_X$, es decir, debemos probar que $\mathbb{P}_Y(f^{-1}(A)) = 0$ para cada $A \in \mathcal{F}_X$ tal que $\mathbb{P}_X(A) = 0$.

  Sea $A \in \mathcal{F}_X$, tal que $\mathbb{P}_X(A) = 0$, notemos que $A \in \mathbb{P}_X^{-1}(0)$, de donde, dado que f es nula-conservadora, se sigue que $f^{-1}(A) \in \mathbb{P}_Y^{-1}(0)$, con esto, por definición de imagen inversa, se sigue que

$$\mathbb{P}_Y(f^{-1}(A)) = 0.$$

- Para la otra implicación, supongamos que para todo conjunto $A \in \mathcal{F}_X$ tal que $\mathbb{P}_X(A) = 0$ se tiene que $\mathbb{P}_Y(f^{-1}(A)) = 0$.

  Sea $A \in \mathbb{P}_X^{-1}(0)$, vamos a demostrar que $f^{-1}(A) \in \mathbb{P}_Y^{-1}(0)$. Tenemos que

  $$\mathbb{P}_X(A) = 0, \ A \in \mathcal{F}_X,$$

  de donde, junto con la hipótesis, obtenemos que

  $$\mathbb{P}_Y(f^{-1}(A)) = 0,$$

  con lo cual $f^{-1}(A) \in \mathbb{P}_Y^{-1}(0)$.

$\square$

PROPOSICIÓN 2.6. *Sean $\overline{X}$, $\overline{Y}$, $\overline{Z}$ espacios de probabilidad y $f\colon \overline{Y} \to \overline{X}$, $g\colon \overline{Z} \to \overline{Y}$ funciones nulas-conservadoras. Se tiene que*

$$f \circ g\colon \overline{Z} \to \overline{X},$$

es nula-conservadora.

Demostración. Gracias a la proposición 2.5 basta con probar que $\mathbb{P}_Z \circ (f \circ g)^{-1} \ll \mathbb{P}_X$.

Sea $A \in \mathcal{F}_X$ tal que $\mathbb{P}_X(A) = 0$, vamos a demostrar que $\mathbb{P}_Z\left((f \circ g)^{-1}(A)\right) = 0$.

Usando el hecho que f es nula conservadora y que $\mathbb{P}_X(A) = 0$, tenemos que

$$\mathbb{P}_Y(f^{-1}(A)) = 0. \tag{2.26}$$

Ahora, como f es $\mathcal{F}_Y/\mathcal{F}_X-$medible tenemos que $f^{-1}(A) \in \mathcal{F}_Y$, de donde, dado que g es nula-conservadora, colegimos que

$$\mathbb{P}_Z\left(g^{-1}\left(f^{-1}(A)\right)\right) = 0,$$

con esto, dado que $(f \circ g)^{-1}(A) = g^{-1}\left(f^{-1}(A)\right)$ para todo $A \in \mathcal{F}_X$ se sigue el resultado. $\square$

Ahora, presentamos la definición de la categoría **Prob**:

DEFINICIÓN 2.24 (Categoría **Prob**). *La categoría* **Prob** *está dada por la siguiente información:*

- ***Objetos***: *Espacios de Probabilidad;*
- ***Flechas***: *Sean $\overline{X}$ y $\overline{Y}$ espacios de probabilidad, el conjunto de flechas entre estos espacios está definido por*

$$\textbf{Prob}(\overline{X}, \overline{Y}) := \{f^{-} | f : \overline{Y} \to \overline{X} \text{ es una función nula-conservadora}\},$$

  donde f^{-} es un símbolo que corresponde únicamente a una función f.

- ***Composición:*** *Sean $f^{-} \colon \overline{X} \to \overline{Y}$ y $g^{-} \colon \overline{Y} \to \overline{Z}$ dos flechas en* **Prob***, la composición viene dada por*

$$g^{-} \circ_{\textbf{Prob}} f^{-} = (f \circ g)^{-},$$

  donde $\circ$ es la composición usual de funciones.

- ***Identidades:*** *Consideraremos la notación Id_X para representar a la función identidad medible de $\overline{X}$ a $\overline{X}$, mientras que escribiremos id_X para una función de identidad de X a X. Así, la flecha identidad de un objeto $\overline{X}$ en* **Prob** *es $Id_{\overline{X}}^{-}$.*

OBSERVACIÓN. Sin riesgo de confusión emplearemos la notación $\circ$ para referirnos tanto a la composición en **Prob** como a la composición en **Set**.

OBSERVACIÓN. **Prob** es una categoría.

Demostración. Gracias a la proposición 2.6, para probar que la Categoría **Prob** está bien definida basta con demostrar que la composición en **Prob** es asociativa y se satisfacen los axiomas de identidad.

1. Para la composición, sean $f^{-} \colon \overline{X} \to \overline{Y}$, $g \colon \overline{Y} \to \overline{Z}$ y $h \colon \overline{Z} \to \overline{W}$ flechas en **Prob**, vamos a demostrar que

$$(h^{-} \circ g^{-}) \circ f^{-} = h^{-} \circ (g^{-} \circ f^{-}).$$

El resultado se sigue usando el hecho de que la asociatividad se satisface en **Set**. En efecto, tenemos que

$$\begin{aligned}(h^- \circ g^-) \circ f^- &= (g \circ h)^- \circ f^-,\\ &= (f \circ (g \circ h))^-,\\ &= ((f \circ g) \circ h)^-,\\ &= h^- \circ (f \circ g)^-,\\ &= h^- \circ (g^- \circ f^-).\end{aligned}$$

2. Para las leyes de identidad, sean $f^- \colon \overline{X} \to \overline{Y}$ y $g^- \colon \overline{Y} \to \overline{X}$ dos flechas en **Prob**, vamos a demostrar que

$$f^- \circ Id_X^- = f^- \quad \text{y} \quad Id_X^- \circ g^- = g^-.$$

Por definición de composición en **Prob** y dado que Id_X es la flecha identidad en **Set**, se tiene que

$$f^- \circ Id_X^- = (Id_X \circ f)^- = f^-.$$

Análogamente se prueba que $Id_X^- \circ g^- = g^-$. □

Motivados por el hecho de que cualquier conjunto unitario es el único objeto inicial de la categoría **Set**, consideremos la siguiente proposición:

PROPOSICIÓN 2.7. *El espacio de probabilidad* $\mathbb{0} := (\{*\}, \{\{*\}, \emptyset\}, \mathbb{P}_*)$*, donde* $\mathbb{P}_*(\{*\}) = 1$ *y* $\mathbb{P}_*(\emptyset) = 0$ *es un objeto inicial de la categoría* ***Prob****.*

Demostración. Sea $\overline{X} \in$ ***Ob***(**Prob**) un objeto en **Prob**, vamos a demostrar que existe una única flecha $!^- \colon \mathbb{0} \to \overline{X}$, es decir, debemos probar que existe una única función $! \colon \overline{X} \to \mathbb{0}$ nula-conservadora.

Sabemos que para cualquier conjunto X existe una única función definida por

$$\begin{aligned} ! \colon X &\longrightarrow \{*\}\\ x &\longmapsto !(x) = *.\end{aligned}$$

Así, basta verificar que $! \colon \overline{X} \to \mathbb{0}$ es nula-conservadora, para ello, debemos probar que $!$ es una función $\mathcal{F}_X/\{\{*\}, \emptyset\}$–medible y $\mathbb{P}_X \circ !^{-1} \ll \mathbb{P}_*$.

1. Sea $A \in \{\{*\}, \emptyset\}$, consideremos los siguientes casos:

   - Si $A = \emptyset$, entonces $!^{-1}(A) = \emptyset \in \mathcal{F}_X$.
   - Si $A = \{*\}$, entonces $!^{-1}(A) = \{B \in \mathcal{F}_X |\, !(B) = \{*\}\} = X \in \mathcal{F}_X$.

   Por lo tanto, $! \colon \overline{X} \to \mathbb{0}$ es una función $\mathcal{F}_X / \{\{*\}, \emptyset\}-$medible.

2. Notemos que el único elemento de medida nula en $\mathbb{0}$ es $\emptyset$. Así, basta con demostrar que $\mathbb{P}_X\left(!^{-1}(\emptyset)\right) = 0$. En efecto, el resultado se sigue dado que $\mathbb{P}_X$ es una medida de probabilidad:

$$\mathbb{P}_X\left(!^{-1}(\emptyset)\right) = \mathbb{P}_X(\emptyset) = 0.$$

   Con esto, hemos probado que $\mathbb{P}_X \circ !^{-1}$ es absolutamente continua con respecto a $\mathbb{P}_*$.

Junto con 1 y 2, podemos concluir que $!^{-} \colon \mathbb{0} \to \overline{X}$ es una flecha en **Prob**, por lo tanto, $\mathbb{0}$ es un objeto inicial de la categoría **Prob**. □

Para terminar esta sección estudiaremos la relación entre la Categoría **Prob** y la Categoría **CPS**. La siguiente proposición establece que $\mathbf{CPS^{op}}$ es una subcategoría de **Prob**:

PROPOSICIÓN 2.8. *Toda función acotada $f \colon \overline{Y} \to \overline{X}$, en el sentido de la definición 2.5, es nula conservadora.*

Demostración. Sea $f \colon \overline{Y} \to \overline{X}$ una función acotada, es decir, existe $M > 0$ tal que

$$\mathbb{P}_Y\left(f^{-1}(A)\right) \leq M\mathbb{P}_X(A), \tag{2.27}$$

para cada $A \in \mathcal{F}_X$. Vamos a demostrar que $\mathbb{P}_Y \circ f^{-1} \ll \mathbb{P}_X$, para ello, sea $A \in \mathcal{F}_X$, tal que $\mathbb{P}_X(A) = 0$, en particular, de (2.27), se tiene que

$$0 \leq \mathbb{P}_Y\left(f^{-1}(A)\right) \leq M\mathbb{P}_X(A) = 0,$$

de donde, podemos concluir que

$$\mathbb{P}_Y\left(f^{-1}(A)\right) = 0,$$

es decir, hemos probado que

$$\mathbb{P}_Y \circ f^{-1} \ll \mathbb{P}_X,$$

como queríamos. $\square$

2.2.2. Funtores en la categoría Prob

En esta sección nos concentraremos en presentar funtores en la categoría **Prob**. Para esto, consideraremos a $\mathcal{L}^\infty(\overline{X})$ como el espacio vectorial de las variables aleatorias real valuadas v tales que $\mathbb{P}_X - \text{ess}\sup_{x\in X} |v(x)| < +\infty$ y a $\mathcal{L}^1(\overline{X})$ como el espacio vectorial de variables aleatorias real valuadas v tales que $\int_X |v|\, d\mathbb{P}_X < +\infty$.

Además, $L^\infty(\overline{X})$ y $L^1(\overline{X})$ representarán a los espacios cocientes de $\mathcal{L}^\infty(\overline{X})/\sim_{\mathbb{P}_X}$ y $\mathcal{L}^1(\overline{X})/\sim_{\mathbb{P}_Y}$, respectivamente, donde

$$\mu_1 \sim_{\mathbb{P}_x} \mu_2 \;\; \text{si} \;\; \mathbb{P}_X\left(\{\omega \,|\, \mu_1(\omega) \neq \mu_2(\omega)\}\right) = 0.$$

DEFINICIÓN 2.25 (Functor **L**). *Sean $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ y $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ dos espacios de probabilidad. El functor $\boldsymbol{L}$: **Prob** $\to$ **Set** está definido por:*

1. *Para objetos:*
$$\begin{aligned} \boldsymbol{L}\colon \boldsymbol{Ob}(\boldsymbol{Prob}) &\longrightarrow \boldsymbol{Ob}(\boldsymbol{Set}) \\ \overline{X} \qquad &\longmapsto \boldsymbol{L}(\overline{X}) := L^\infty(\overline{X}). \end{aligned}$$

2. *Para flechas:*
$$\begin{aligned} \boldsymbol{L}\colon \boldsymbol{Prob}(\overline{X}, \overline{Y}) &\longrightarrow \boldsymbol{Set}(\boldsymbol{L}(\overline{X}), \boldsymbol{L}(\overline{Y})) \\ [f^- \colon \overline{X} \to \overline{Y}] &\longmapsto [\boldsymbol{L}(f^-) : L^\infty(\overline{X}) \to L^\infty(\overline{Y})], \end{aligned}$$

   donde, $\boldsymbol{L}(f^-)$ es una función tal que

$$\begin{aligned} \boldsymbol{L}(f^-) : L^\infty(\overline{X}) &\longrightarrow L^\infty(\overline{Y}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto [u \circ f]_{\sim\mathbb{P}_Y} \end{aligned}$$

OBSERVACIÓN. El functor **L**: **Prob** $\to$ **Set** está bien definido.

Demostración. La demostración la haremos en dos pasos, primero probaremos que $\mathbf{L}(f^-)$ está bien definida. Luego, probaremos que el functor **L** preserva la composición y las identidades.

1. Notemos que el functor $\mathbf{L}$ está bien definido sobre objetos, pues envía un espacio de probabilidad $\overline{X}$ al conjunto $L^{\infty}(\overline{X})$.

   Así, basta con demostrar que la función $\mathbf{L}(f^{-})$ está bien definida. Para ello, sean $u_1, u_2 \in \mathcal{L}^{\infty}(\overline{X})$ tales que $u_1 \sim_{\mathbb{P}_X} u_2$, vamos a demostrar que $u_1 \circ f \sim_{\mathbb{P}_Y} u_2 \circ f$, para lo cual, debemos probar que

   $$\mathbb{P}_Y\left(\{\omega \in Y : (u_1 \circ f)(\omega) \neq (u_2 \circ f)(\omega)\}\right) = 0.$$

   Puesto que $u_1 \sim_{\mathbb{P}_X} u_2$, sabemos que $\mathbb{P}_X\left(\{\omega \in X : u_1(\omega) \neq u_2(\omega)\}\right) = 0$, de donde, dado que $\mathbb{P}_Y \circ f^{-1} \ll \mathbb{P}_X$, tenemos que

   $$\begin{aligned} \mathbb{P}_Y\left(f^{-1}\left(\{\omega \in X : u_1(\omega) \neq u_2(\omega)\}\right)\right) &= (\mathbb{P}_Y \circ f^{-1})\left(\{\omega \in X : u_1(\omega) \neq u_2(\omega)\}\right), \\ &= 0. \end{aligned} \tag{2.28}$$

   Ahora, probemos que

   $$\{\omega \in Y : (u_1 \circ f)(\omega) \neq (u_2 \circ f)(\omega)\} \subset f^{-1}\left(\{\omega \in X : u_1(\omega) \neq u_2(\omega)\}\right),$$

   lo cual es equivalente a demostrar que

   $$f^{-1}\left(\{\omega \in X : u_1(\omega) = u_2(\omega)\}\right) \subset \{\omega \in Y : (u_1 \circ f)(\omega) = (u_2 \circ f)(\omega)\}. \tag{2.29}$$

   Sea $x \in f^{-1}\left(\{\omega \in X : u_1(\omega) = u_2(\omega)\}\right)$, tenemos que $f(x) \in \{\omega \in X : u_1(\omega) = u_2(\omega)\}$, es decir,

   $$u_1(f(x)) = u_2(f(x)),$$

   con lo cual $x \in \{\omega \in Y : (u_1 \circ f)(\omega) = (u_2 \circ f)(\omega)\}$, por lo tanto, hemos probado la contenencia (2.29).

   Finalmente, por la monotonía de $\mathbb{P}_Y$ y junto con (2.28), obtenemos que

   $$\begin{aligned} \mathbb{P}_Y\left(\{\omega \in Y : (u_1 \circ f)(\omega) \neq (u_2 \circ f)(\omega)\}\right) &\leq \mathbb{P}_Y\left(f^{-1}\left(\{\omega \in X : u_1(\omega) \neq u_2(\omega)\}\right)\right), \\ &= 0, \end{aligned}$$

así, podemos concluir que

$$\mathbb{P}_Y\left(\{\omega \in Y : (u_1 \circ f)(\omega) \neq (u_2 \circ f)(\omega)\}\right) = 0,$$

como queríamos.

2. Ahora, vamos a demostrar que **L** preserva la composición y las identidades.

   Sean $f^-\colon \overline{X} \to \overline{Y}$ y $g^-\colon \overline{Y} \to \overline{Z}$ dos flechas en **Prob**, debemos probar que

$$\mathbf{L}(g^- \circ f^-) = \mathbf{L}(g^-) \circ \mathbf{L}(f^-).$$

   Por un lado, aplicando la definición de **L** para flechas, tenemos que

$$\begin{aligned} \mathbf{L}\left(g^- \circ f^-\right) : L^\infty(\overline{X}) &\longrightarrow L^\infty(\overline{Z}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto [u \circ (f \circ g)]_{\sim\mathbb{P}_Z}. \end{aligned} \tag{2.30}$$

   Por otro lado, se tiene que

$$\begin{aligned} \mathbf{L}\left(f^-\right) : L^\infty(\overline{X}) &\longrightarrow L^\infty(\overline{Y}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto [u \circ f]_{\sim\mathbb{P}_Y}. \end{aligned} \quad \text{y} \quad \begin{aligned} \mathbf{L}\left(g^-\right) : L^\infty(\overline{Y}) &\longrightarrow L^\infty(\overline{Z}) \\ [w]_{\sim\mathbb{P}_Y} &\longmapsto [w \circ g]_{\sim\mathbb{P}_Z}, \end{aligned}$$

   de donde,

$$\begin{aligned} \mathbf{L}\left(g^-\right) \circ \mathbf{L}\left(f^-\right) : L^\infty(\overline{X}) &\longrightarrow L^\infty(\overline{Z}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto \left(\mathbf{L}\left(g^-\right) \circ \mathbf{L}\left(f^-\right)\right)\left([u]_{\sim\mathbb{P}_X}\right), \end{aligned}$$

   con esto, para todo $u \in \mathcal{L}(\overline{X})$, obtenemos que

$$\left(\mathbf{L}\left(g^-\right) \circ \mathbf{L}\left(f^-\right)\right)\left([u]_{\sim\mathbb{P}_X}\right) = \mathbf{L}(g^-)\left([u \circ f]_{\sim\mathbb{P}_y}\right) = [u \circ (f \circ g)]_{\sim\mathbb{P}_Z},$$

   luego, junto con (2.30), podemos concluir que

$$\mathbf{L}(g^- \circ f^-) = \mathbf{L}(g^-) \circ \mathbf{L}(f^-).$$

   Finalmente, consideremos $Id_X^-\colon \overline{X} \to \overline{X}$ la flecha identidad en **Prob**, vamos a demostrar que

$$\mathbf{L}(Id_X^-) = Id,$$

donde
$$\begin{aligned} Id : L^{\infty}(\overline{X}) &\longrightarrow L^{\infty}(\overline{X}) \\ [u]_{\sim \mathbb{P}_X} &\longmapsto [u]_{\sim \mathbb{P}_X}. \end{aligned}$$

Aplicando la definición de **L** para flechas, tenemos que

$$\begin{aligned} \mathbf{L}\left(Id_{\overline{X}}\right) : L^{\infty}(\overline{X}) &\longrightarrow L^{\infty}(\overline{X}) \\ [u]_{\sim \mathbb{P}_X} &\longmapsto [u \circ Id_X]_{\sim \mathbb{P}_x} = [u]_{\sim \mathbb{P}_x}, \end{aligned}$$

con lo cual se sigue el resultado. □

Para presentar el siguiente funtor es necesario recordar la definición de completación de un espacio de probabilidad.

DEFINICIÓN 2.26. *Sea $(X, \mathcal{F}_X, \mathbb{P}_X)$ un espacio de probabilidad. Consideremos*

1. $\mathcal{F}_X^* := \{F \subset X : \exists A, B \in \mathcal{F}_X, A \subset F \subset B$ **y** $\mathbb{P}_X(B \setminus A) = 0\}$;
2. *Para $F \in \mathcal{F}_X^*$, se define $\mathbb{P}_X^*$ por $\mathbb{P}_X^*(E) := \mathbb{P}_X(F) = \mathbb{P}_X(B)$ donde $A, B \in \mathbb{F}_X$ satisfacen $A \subset F \subset B$* **y** *$\mathbb{P}_X(B \setminus A) = 0$.*

Se tiene que $(X, \mathcal{F}_X^*, \mathbb{P}_X^*)$ es un espacio de probabilidad y se denomina la completación *de $(X, \mathcal{F}_X, \mathbb{P}_X)$.*

OBSERVACIÓN. Si consideramos el conjunto $\mathcal{N} = \{N \in \mathcal{F}_X : \exists Z \in \mathcal{F}_X$ tal que $N \subset Z$ y $\mathbb{P}_X(Z) = 0\}$, entonces
$$\mathcal{F}_x^* = \sigma(\mathcal{F}_X, \mathcal{N}).$$

Así, podemos considerar la siguiente definición:

DEFINICIÓN 2.27 (Funtor $\mathcal{C}$). *Sean $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$* **y** *$\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ dos espacios de probabilidad. El funtor $\mathcal{C}$: **Prob** $\to$ **Prob** está definido por:*

1. *Para objetos:*

$$\begin{aligned} \mathcal{C} : \textbf{\textit{Ob}}(\textbf{\textit{Prob}}) &\longrightarrow \textbf{\textit{Ob}}(\textbf{\textit{Prob}}) \\ (X, \mathcal{F}_X, \mathbb{P}_X) &\longmapsto \mathcal{C}\left((X, \mathcal{F}_X, \mathbb{P}_X)\right) := (X, \mathcal{F}_X^*, \mathbb{P}_X^*), \end{aligned}$$

donde $(X, \mathcal{F}_X^*, \mathbb{P}_X^*)$ es la completación de $(X, \mathcal{F}_X, \mathbb{P}_X)$.

2. *Para flechas:*

$$\begin{aligned} \mathcal{C}:\textbf{\textit{Prob}}(\overline{X},\overline{Y}) &\longrightarrow \textbf{\textit{Prob}}(\mathcal{C}(\overline{X}),\mathcal{C}(\overline{Y})) \\ f^{-}\colon \overline{X}\to\overline{Y} &\longmapsto f^{-}\colon (X,\mathcal{F}_X^*,\mathbb{P}_X^*)\to(Y,\mathcal{F}_Y^*,\mathbb{P}_Y^*). \end{aligned}$$

El funtor $\mathcal{C}$ *se denomina* funtor de completación.

Observación. El funtor $\mathcal{C}\colon$ **Prob** $\to$ **Prob** está bien definido.

Demostración. Notemos que el funtor $\mathcal{C}$ envía espacios de probabilidad en espacios de probabilidad. Así, basta probar que $f^{-}\colon (X,\mathcal{F}_X^*,\mathbb{P}_X^*)\to (Y,\mathcal{F}_Y^*,\mathbb{P}_Y^*)$ es una flecha en **Prob**. Para lo cual, debemos demostrar que:

1. La función $f\colon (Y,\mathcal{F}_Y)\to(X,\mathcal{F}_X)$ es $\mathcal{F}_Y^*/\mathcal{F}_X^*-$medible.

   Sea $F\in\mathcal{F}_X^*$, vamos a demostrar que $f^{-1}(F)\in\mathcal{F}_Y^*$, por definición de completación, debemos probar que existen C ,$D\in\mathcal{F}_Y$ tales que

   $$C\subset f^{-1}(F)\subset D \quad \text{y} \quad \mathbb{P}_Y(D\setminus C)=0.$$

   Por definición de completación aplicada a $F\in\mathcal{F}_X^*$, sabemos existen $A,B\in\mathcal{F}_X$ tales que

   $$A\subset F\subset B \quad \text{y} \quad \mathbb{P}_X(B\setminus A)=0,$$

   de donde, dado que $f\colon\overline{Y}\to\overline{X}$ es nula-conservadora, obtenemos que

   $$f^{-1}(A)\subset f^{-1}(F)\subset f^{-1}(B) \quad \text{y} \quad \mathbb{P}_Y(f^{-1}(B)\setminus f^{-1}(A))=\mathbb{P}_Y(f^{-1}(B\setminus A))=0.$$

   Con esto, dado que $f\colon(Y,\mathcal{F}_Y)\to(X,\mathcal{F}_X)$ es $\mathcal{F}_Y/\mathcal{F}_X-$medible, basta tomar $C=f^{-1}(A)\in\mathcal{F}_Y$ y $D=f^{-1}(B)\in\mathcal{F}_Y$ y se sigue el resultado.

2. $\mathbb{P}_Y^*\circ f^{-1}\ll\mathbb{P}_X^*$.

   Sea $F\in\mathcal{F}_X^*$ tal que $\mathbb{P}_X^*(F)=0$, por definición de completación, sabemos que existe $B\in\mathcal{F}_X$ tal que

   $$F\subset B \quad \text{y} \quad \mathbb{P}_X^*(F)=\mathbb{P}_X(B)=0.$$

   Ahora, dado que $\mathbb{P}_Y\circ f^{-1}\ll\mathbb{P}_X$, en particular, tenemos que $\mathbb{P}_Y\left(f^{-1}(B)\right)=$

0. Así, obtenemos que

$$0 \leq \mathbb{P}_Y^* \left(f^{-1}(F)\right) \leq \mathbb{P}_Y^* \left(f^{-1}(B)\right) = \mathbb{P}_Y \left(f^{-1}(B)\right) = 0,$$

es decir,

$$\mathbb{P}_Y^* \left(f^{-1}(F)\right) = 0. \qquad \square$$

2.2.3. Funtor esperanza condicional

A continuación, presentamos algunos resultados necesarios para definir el funtor más importante de esta sección.

OBSERVACIÓN. Sean $f^- \colon \overline{X} \to \overline{Y}$ una flecha en **Prob** y $v \in \mathcal{L}^1(\overline{Y})$. Para cada $B \in \mathcal{F}_Y$, consideremos

$$v^*(B) = \int_B v \, d\mathbb{P}_Y.$$

Sabemos que la expresión precedente determina una medida con signo sobre $(Y, \mathcal{F}_Y)$. Así, consideremos la medida imagen de v^* con respecto a f, definida por

$$v^* \circ f^{-1}(A) = \int_{f^{-1}(A)} v \, d\mathbb{P}_Y,$$

para cada $A \in \mathcal{F}_X$, la cual es una medida con signo definida sobre $(X, \mathcal{F}_X)$ y absolutamente continua con respecto a $\mathbb{P}_X$.

Demostración. Vamos a demostrar que $v^* \circ f^{-1}$ es absolutamente continua con respecto a $\mathbb{P}_X$, para ello, sea $A \in \mathcal{F}_X$ tal que $\mathbb{P}_X(A) = 0$, debemos probar que $(v^* \circ f^{-1})(A) = 0$.

Dado que $f \colon \overline{Y} \to \overline{X}$ es nula-conservadora y $A \in \mathcal{F}_X$ es un conjunto de medida nula, tenemos que $\mathbb{P}_Y \left(f^{-1}(A)\right) = 0$, con lo cual

$$v^* \circ f^{-1}(A) = \int_{f^{-1}(A)} v \, d\mathbb{P}_Y = \int_Y v \cdot \mathbb{1}_{f^{-1}(A)} \, d\mathbb{P}_Y. \tag{2.31}$$

Ahora, notemos que $\varphi := v \cdot \mathbb{1}_{f^{-1}(A)} = 0$ $\mathbb{P}_Y-$casi todas partes. En efecto, consideremos el conjunto $N = \{\omega \in Y : \varphi(\omega) \neq 0\}$ y probemos que $\mathbb{P}_Y(N) = 0$. Para esto, demostremos que $N \subset f^{-1}(A)$, lo cual es equivalente a tener que $f^{-1}(A^C) \subset N^C$.

Sea $x \in f^{-1}(A^C) = (f^{-1}(A))^C$, tenemos que

$$\mathbb{1}_{f^{-1}(A)}(x) = 0,$$

de donde $\varphi(x) = v(x) \cdot \mathbb{1}_{f^{-1}(A)}(x) = 0$, con lo cual $x \in N^C = \{\omega \in Y : \varphi(\omega) = 0\}$.

Así, por la monotonía de $\mathbb{P}_Y$, se sigue que

$$0 \leq \mathbb{P}_Y(N) \leq \mathbb{P}_Y\left(f^{-1}(A)\right) = 0,$$

es decir, $\mathbb{P}_Y(N) = 0$, por lo tanto, se sigue el resultado.

Con esto, tenemos que

$$\int_Y v \cdot \mathbb{1}_{f^{-1}(A)} \, d\mathbb{P}_Y = \int_Y 0 \, d\mathbb{P}_Y = 0,$$

de donde, junto con (2.31), podemos concluir que

$$(v^* \circ f^{-1})(A) = 0. \qquad \square$$

Gracias a la observación 2.2.3 tenemos la siguiente definición:

Definición 2.28 (Esperanza condicional de v por f^-)**.** *Sean $\overline{X}$ y $\overline{Y}$ espacios de probabilidad, $f^- \colon \overline{X} \to \overline{Y}$ una flecha en* ***Prob*** *y $v \in \mathcal{L}^1(\overline{Y})$. Por el teorema de Radon-Nikodym aplicado a la medida $\mathbb{P}_X$ y a la medida con signo $v^* \circ f^{-1}$ absolutamente continua con respecto a $\mathbb{P}_X$, sabemos existe un único elemento ($\mathbb{P}_X-$casi seguramente) denotado por $E^{f^-}(v) \in \mathcal{L}^1(\overline{X})$ tal que*

$$\int_A E^{f^-}(v) \, d\mathbb{P}_X = \int_{f^{-1}(A)} v \, d\mathbb{P}_Y,$$

para cada $A \in \mathcal{F}_X$. Llamaremos al elemento $E^{f^-}(v)$ la esperanza condicional de v por f^-.

Previo a la presentación del funtor esperanza condicional, expondremos dos proposiciones que nos garantizan la correcta definición del mismo.

Proposición 2.9. *Sea $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ un espacio de probabilidad, para $u \in \mathcal{L}^1(\overline{X})$, se tiene que*

$$E^{Id_{\overline{X}}^-}(u) \sim_{\mathbb{P}_X} u.$$

Demostración. Sean $u \in \mathcal{L}^1(\overline{X})$ y $A \in \mathcal{F}_X$, dado que $Id_X^{-1}(A) = A$, la esperanza condicional de u por Id_X^- satisface que

$$\int_A E^{Id_X^-}(u)\, d\mathbb{P}_X = \int_{Id^{-1}(A)} u\, d\mathbb{P}_X, = \int_A u\, d\mathbb{P}_X,$$

para todo $A \in \mathcal{F}_X$, de donde,

$$E^{Id_X^-}(u) = u \ \ \mathbb{P}_X - \text{casi seguramente},$$

lo cual es equivalente a

$$E^{Id_X^-}(u) \sim_{\mathbb{P}_X} u.$$

□

PROPOSICIÓN 2.10. *Sean $f\colon \overline{X} \to \overline{Y}$ y $g\colon \overline{Y} \to \overline{Z}$ flechas en* ***Prob****, se tiene que*

1. *Para $v_1, v_2 \in \mathcal{L}(\overline{Y})$, si $v_1 \sim_{\mathcal{P}_Y} v_2$, entonces $E^{f^-}(v_1) \sim_{\mathbb{P}_X} E^{f^-}(v_2)$;*
2. *Para $w \in \mathcal{L}(\overline{Z})$, se tiene que $E^{f^-}\left(E^{g^-}(w)\right) \sim_{\mathbb{P}_X} E^{g^- \circ f^-}(w)$.*

Demostración. 1. Sean $v_1, v_2 \in \mathcal{L}(\overline{Y})$ tales que $v_1 \sim_{\mathbb{P}_Y} v_2$, dado que $v_1 = v_2$ $\mathbb{P}_Y-$casi seguramente, tenemos que

$$v_1^*(B) = \int_B v_1\, d\mathbb{P}_Y = \int_B v_2\, d\mathbb{P}_Y = v_2^*(B),$$

para cada $B \in \mathcal{F}_Y$. Así, si consideramos $f^-\colon \overline{X} \to \overline{Y}$ una flecha en **Prob**, tenemos que

$$v_1^* \circ f^{-1}(A) = \int_{f^{-1}(A)} v_1\, d\mathbb{P}_Y = \int_{f^{-1}(A)} v_2\, d\mathbb{P}_Y = v_2^* \circ f^{-1}(A),$$

para cada $A \in \mathcal{F}_X$, de donde, dado que

$$\int_A E^{f^-}(v_1)\, d\mathbb{P}_X = \int_{f^{-1}(A)} v_1\, d\mathbb{P}_Y \quad \text{y} \quad \int_A E^{f^-}(v_2)\, d\mathbb{P}_X = \int_{f^{-1}(A)} v_2\, d\mathbb{P}_Y,$$

obtenemos que

$$\int_A E^{f^-}(v_1)\, d\mathbb{P}_X = \int_A E^{f^-}(v_2)\, d\mathbb{P}_X,$$

con esto, se tiene que

$$E^{f^-}(v_1) = E^{f^-}(v_2), \ \ \mathbb{P}_X - \text{casi-seguramente},$$

lo cual es equivalente a tener que

$$E^{f^-}(v_1) \sim_{\mathbb{P}_X} E^{f^-}(v_2).$$

2. Sean $f\colon \overline{X} \to \overline{Y}$, $g\colon \overline{Y} \to \overline{Z}$ flechas en **Prob** y $w \in \mathcal{L}(\overline{Z})$.

   En primer lugar, sabemos que la esperanza condicional de w por g^-, denotada por $E^{g^-}(w) \in \mathcal{L}^1(\overline{Y})$, es tal que

$$\int_B E^{g^-}(w)\, d\mathbb{P}_Y = \int_{g^{-1}(B)} w\, d\mathbb{P}_Z, \tag{2.32}$$

   para cada $B \in \mathcal{F}_Y$.

   Luego, la esperanza condicional de $E^{g^-}(w)$ por f^- satisface que

$$\int_A E^{f^-}\left(E^{g^-}(w)\right)\, d\mathbb{P}_X = \int_{f^{-1}(A)} E^{g^-}(w)\, d\mathbb{P}_Y, \tag{2.33}$$

   para cada $A \in \mathbb{P}_X$.

   Ahora, dado que $f^{-1}(A) \in \mathcal{F}_Y$, en particular de (2.32), tenemos

$$\int_{f^{-1}(A)} E^{g^-}(w)\, d\mathbb{P}_Y = \int_{g^{-1}(f^{-1}(A))} w\, d\mathbb{P}_Z,$$

   de donde, junto con (2.33), se sigue que

$$\int_A E^{f^-}\left(E^{g^-}(w)\right)\, d\mathbb{P}_X = \int_{(f\circ g)^{-1}(A)} w\, d\mathbb{P}_Z. \tag{2.34}$$

   Finalmente, la esperanza condicional de w por $g^- \circ f^-\colon \overline{X} \to \overline{Z}$, satisface que

$$\int_A E^{g^-\circ f^-}(w)\, d\mathbb{P}_X = \int_{(f\circ g)^{-1}(A)} w\, d\mathbb{P}_Z,$$

   para cada $A \in \mathcal{F}_X$. Combinando la identidad precedente con (2.34), obtenemos que

$$\int_A E^{f^-}\left(E^{g^-}(w)\right)\, d\mathbb{P}_X = \int_A E^{g^-\circ f^-}(w)\, d\mathbb{P}_X,$$

con lo cual, podemos concluir que

$$E^{f^-}\left(E^{g^-}(w)\right) \sim_{\mathbb{P}_X} E^{g^-\circ f^-}(w),$$

como queríamos. $\square$

Tenemos las herramientas necesarias para presentar el functor esperanza condicional:

Definición 2.29 (Funtor esperanza condicional ε)**.** *Sean $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ y $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ dos espacios de probabilidad. El functor $\varepsilon\colon$ **Prob**^{op} $\to$ **Set** está definido por:*

1. *Para objetos:*

$$\begin{aligned} \varepsilon : \boldsymbol{Ob}(\boldsymbol{Prob^{op}}) &\longrightarrow \boldsymbol{Ob}(\mathbf{Set}) \\ \overline{X} &\longmapsto \varepsilon(\overline{X}) := L^1(\overline{X}). \end{aligned}$$

2. *Para flechas:*

$$\begin{aligned} \varepsilon : \boldsymbol{Prob^{op}}(\overline{Y}, \overline{X}) &\longrightarrow \mathbf{Set}(\varepsilon(\overline{Y}), \varepsilon(\overline{X})) \\ f^-\colon \overline{X} \to \overline{Y} &\longmapsto \varepsilon\left(f^-\right) : L^1(\overline{Y}) \to L^1(\overline{X}), \end{aligned}$$

donde, $\varepsilon\left(f^-\right)$ es una función tal que

$$\begin{aligned} \varepsilon\left(f^-\right) : L^1(\overline{Y}) &\longrightarrow L^1(\overline{X}) \\ [v]_{\sim\mathbb{P}_Y} &\longmapsto \left[E^{f^-}(v)\right]_{\sim\mathbb{P}_X}. \end{aligned}$$

El funtor ε se denomina funtor esperanza condicional.

Observación. La proposición 2.9 nos garantiza que el functor ε preserva las identidades. Consideremos $Id_{\overline{X}}^-\colon \overline{X} \to \overline{X}$ la flecha identidad en **Prob**, aplicando la definición de ε para flechas, tenemos que

$$\begin{aligned} \varepsilon\left(Id_{\overline{X}}^-\right) : L^1(\overline{X}) &\longrightarrow L^1(\overline{X}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto \left[E^{f^-}(u)\right]_{\sim\mathbb{P}_X} = [u]_{\sim\mathbb{P}_X}. \end{aligned}$$

Por un lado, la primera parte de la proposición 2.10 nos permite probar que la función $\varepsilon\left(f^-\right)$ está bien definida.

Por otro lado, la segunda parte nos asegura que el functor ε preserva la composición. En efecto, sean $f\colon \overline{X} \to \overline{Y}$ y $g\colon \overline{Y} \to \overline{Z}$ flechas en **Prob**. Calculemos $\boldsymbol{\varepsilon}(g^{-} \circ f^{-})$:

$$\begin{aligned}\boldsymbol{\varepsilon}\left(g^{-} \circ f^{-}\right) : L^1(\overline{Z}) &\longrightarrow L^1(\overline{X})\\ [w]_{\sim \mathbb{P}_Z} &\longmapsto \left[E^{g \circ f^{-}}(w)\right]_{\sim \mathbb{P}_X},\end{aligned}$$

de donde, por la Proposición 2.10, se tiene que

$$\begin{aligned}\boldsymbol{\varepsilon}\left(g^{-} \circ f^{-}\right)\left([w]_{\sim \mathbb{P}_Z}\right) &= \left[E^{g^{-} \circ f^{-}}(w)\right]_{\sim \mathbb{P}_X},\\ &= \left[E^{f^{-}}\left(E^{g^{-}}(w)\right)\right]_{\sim \mathbb{P}_X},\\ &= \left(\boldsymbol{\varepsilon}(f^{-}) \circ \boldsymbol{\varepsilon}(g^{-})\right)\left([w]_{\sim \mathbb{P}_Z}\right),\end{aligned}$$

para cada $w \in \mathcal{L}^1(\overline{Z})$. Con lo cual, podemos concluir que

$$\boldsymbol{\varepsilon}\left(f^{-}_{\text{op}} \circ g^{-}_{\text{op}}\right) = \boldsymbol{\varepsilon}\left(g^{-} \circ f^{-}\right) = \boldsymbol{\varepsilon}(f^{-}) \circ \boldsymbol{\varepsilon}(g^{-}).$$

Propiedades del Functor Esperanza Condicional

En esta sección estudiaremos las propiedades que satisface el functor esperanza condicional $\boldsymbol{\varepsilon}$, análogas a las que satisface la esperanza condicional usual.

PROPOSICIÓN 2.11 (Linealidad). *Sea $f^{-}\colon \overline{X} \to \overline{Y}$ una flecha en **Prob**. Para cualesquier par de variables aleatorias $u, v \in \mathcal{L}^1(\overline{Y})$ y escalares $\alpha, \beta \in \mathbb{R}$, se tiene que*

$$E^{f^{-}}(\alpha u + \beta v) \sim_{\mathbb{P}_X} \alpha E^{f^{-}}(u) + \beta E^{f^{-}}(v).$$

Demostración. Sean $f^{-}\colon \overline{X} \to \overline{Y}$ una flecha en **Prob**, $u, v \in \mathcal{L}^1(\overline{Y})$ variables aleatorias y $\alpha, \beta \in \mathbb{R}$ escalares, sabemos que la esperanza condicional de $\alpha u + \beta v$ por f, satisface que

$$\int_A E^{f^{-}}(\alpha u + \beta v)\, d\mathbb{P}_X = \int_{f^{-1}(A)} (\alpha u + \beta v)\, d\mathbb{P}_Y,$$

para cada $A \in \mathcal{F}_X$, de donde, usando la linealidad de la integral, obtenemos que

$$\int_A E^{f^{-}}(\alpha u + \beta v)\, d\mathbb{P}_X = \int_{f^{-1}(A)} (\alpha u + \beta v)\, d\mathbb{P}_Y$$

$$
\begin{aligned}
&= \int_{f^{-1}(A)} \alpha u \, d\mathbb{P}_Y + \int_{f^{-1}(A)} \beta v \, d\mathbb{P}_Y \\
&= \alpha \int_{f^{-1}(A)} u \, d\mathbb{P}_Y + \beta \int_{f^{-1}(A)} v \, d\mathbb{P}_Y,
\end{aligned}
$$

para cada $A \in \mathcal{F}_X$.

Por otro lado, notemos que

$$\int_A E^{f^-}(u) \, d\mathbb{P}_X = \int_{f^{-1}(A)} u \, d\mathbb{P}_Y \quad \text{y} \quad \int_A E^{f^-}(v) \, d\mathbb{P}_X = \int_{f^{-1}(A)} v \, d\mathbb{P}_Y,$$

para cada $A \in \mathcal{F}_X$, con lo cual se tiene que

$$
\begin{aligned}
\int_A E^{f^-}(\alpha u + \beta v) \, d\mathbb{P}_X &= \alpha \int_A E^{f^-}(u) \, d\mathbb{P}_X + \beta \int_A E^{f^-}(v) \, d\mathbb{P}_X, \\
&= \int_A \Big(\alpha E^{f^-}(u) \, d\mathbb{P}_X + \beta E^{f^-}(v) \Big) \, d\mathbb{P}_X,
\end{aligned}
$$

para cada $A \in \mathcal{F}_X$. Así, podemos concluir que

$$E^{f^-}(\alpha u + \beta v) \, d\mathbb{P}_X = \alpha E^{f^-}(u) \, d\mathbb{P}_X + \beta E^{f^-}(v) \;\; \mathbb{P}_X - \text{casi seguramente},$$

como queríamos. □

PROPOSICIÓN 2.12 (Positividad). *Sean $f^- \colon \overline{X} \to \overline{Y}$ una flecha en **Prob** y $v \in \mathcal{L}^1(\overline{Y})$ una variable aleatoria. Si v es positiva $\mathbb{P}_Y-$ casi seguramente, entonces*

$$E^{f^-}(u) \geq 0 \quad \mathbb{P}_X - \textit{casi seguramente.}$$

Demostración. Si v es positiva $\mathbb{P}_Y-$ casi seguramente, entonces

$$v^* \circ f^{-1}(A) = \int_{f^{-1}(A)} v \, d\mathbb{P}_Y,$$

para cada $A \in \mathcal{F}_X$, es una medida sobre $(X, \mathcal{F}_X)$. De donde, dado que $E^{f^-}(u) \in \mathcal{L}^1(\overline{X})$ es la derivada de Radon Nikodym de $v^* \circ f$, obtenemos que

$$E^{f^-}(u) := \frac{d(v^* \circ f)}{d\mathbb{P}_X} \geq 0 \quad \mathbb{P}_X - \text{casi seguramente.}$$

□

PROPOSICIÓN 2.13 (Monotonía). *Sean $f^- \colon \overline{X} \to \overline{Y}$ una flecha en **Prob***

y $v_1, v_2 \in \mathcal{L}^1(\overline{Y})$ variables aleatorias. Si $v_1 \leq v_2$ $\mathbb{P}_Y-$ casi seguramente, entonces $E^{f^-}(v_1) \leq E^{f^-}(v_2)$ $\mathbb{P}_X-$ casi seguramente,

Demostración. Si $v_1 \leq v_2$ $\mathbb{P}_Y-$ casi seguramente, entonces

$$v_2 - v_1 \geq 0 \quad \mathbb{P}_Y - \text{casi seguramente},$$

de donde, por la proposición 2.12 aplicada a la variable aleatoria $v = v_2 - v_1$, obtenemos que

$$E^{f^-}(v_2 - v_1) \geq 0 \quad \mathbb{P}_X - \text{casi seguramente},$$

con lo cual, junto con la proposición 2.11, podemos concluir que

$$E^{f^-}(v_2) \geq E^{f^-}(v_1) \quad \mathbb{P}_X - \text{casi seguramente},$$

como queríamos. □

Previo a la presentación de la siguiente proposición consideraremos la siguiente notación:

Observación. Si $v \in \mathcal{L}^1(\overline{Y})$ es una variable aleatoria y $(v_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias en $\mathcal{L}^1(\overline{Y})$, emplearemos la notación $v_n \uparrow v$, cuando:

- $v_n \leq v_{n+1}$ para cada $n \in \mathbb{N}$; y,
- $\lim\limits_{n\to\infty} v_n = v$.

Proposición 2.14 (Convergencia Monónota). *Sean $f^- : \overline{X} \to \overline{Y}$ una flecha en **Prob**, $v \in \mathcal{L}^1(\overline{Y})$ una variable aleatoria y $(v_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias en $\mathcal{L}^1(\overline{Y})$. Si $0 \leq v_n \uparrow v$ $\mathbb{P}_Y-$casi seguramente, entonces $0 \leq E^{f^-}(v_n) \uparrow E^{f^-}(v)$ $\mathbb{P}_X-$casi seguramente.*

Demostración. Sea $n \in \mathbb{N}$, dado que $v_n \leq v_{n+1}$ $\mathbb{P}_Y-$casi seguramente, por la proposición 2.13, tenemos que

$$E^{f^-}(v_n) \leq E^{f^-}(v_{n+1}) \quad \mathbb{P}_X - \text{casi seguramente},$$

además, dado que $E^{f^-}(v_n) \in \mathcal{L}^1(\overline{X})$, sabemos existe $h := \limsup_{n\to\infty} E^{f^-}(v_n) \in \mathcal{L}^1(\overline{X})$ tal que

$$E^{f^-}(v_n) \leq h \quad \text{y} \quad \lim_{n\to\infty} E^{f^-}(v_n) = h \qquad \mathbb{P}_X - \text{casi seguramente.} \tag{2.35}$$

Por otro lado, por la proposición 2.12 sabemos que

$$E^{f^-}(v_n) \geq 0 \quad \mathbb{P}_X - \text{casi seguramente,}$$

de donde, junto con (2.35) podemos concluir que

$$0 \leq E^{f^-}(v_n) \uparrow E^{f^-}(v) \quad \mathbb{P}_X - \text{casi seguramente.}$$

Así, basta con demostrar que $h = E^{f^-}(v)$ $\mathbb{P}_X-$casi seguramente. En efecto, dado que se satisfacen las hipótesis del teorema de convergencia monótona, tenemos que

$$\begin{aligned}
\int_A h \, d\mathbb{P}_X &= \limsup_{n\to\infty} \int_A E^{f^-}(v_n) \, d\mathbb{P}_X, \\
&= \limsup_{n\to\infty} \int_{f^{-1}(A)} v_n \, d\mathbb{P}_Y, \\
&= \int_{f^{-1}(A)} \limsup_{n\to\infty} v_n \, d\mathbb{P}_Y,
\end{aligned}$$

para cada $A \in \mathcal{F}_X$, de donde, dado que $(v_n)_{n\in\mathbb{N}}$ es convergente sabemos que $\limsup_{n\to\infty} v_n = \lim_{n\to\infty} v_n = v$. Así, junto con la identidad precedente, se sigue que

$$\begin{aligned}
\int_A h \, d\mathbb{P}_X &= \int_{f^{-1}(A)} \lim_{n\to\infty} v_n \, d\mathbb{P}_Y, \\
&= \int_{f^{-1}(A)} v \, d\mathbb{P}_Y, \\
&= \int_A E^{f^-}(v) \, d\mathbb{P}_X
\end{aligned}$$

para cada $A \in \mathcal{F}_X$. Con lo cual, podemos concluir que

$$h = E^{f^-}(v) \quad \mathbb{P}_X - \text{casi-seguramente.}$$

Con esto, hemos probado que

$$0 \leq E^{f^-}(v_n) \uparrow E^{f^-}(v) \quad \mathbb{P}_X - \text{casi seguramente},$$

como queríamos. □

Observación. Notemos que en la proposición 2.6 exigimos que $v \geq 0$ $\mathbb{P}_Y-$casi seguramente y que $v_n \geq 0$ $\mathbb{P}_Y$-casi seguramente para cada $n \in \mathbb{N}$. Con esto, conseguimos que

$$v^* \circ f^{-1}(A) = \int_{f^{-1}(A)} v\, d\mathbb{P}_Y \quad \text{y} \quad v_n^* \circ f^{-1}(A) = \int_{f^{-1}(A)} v_n\, d\mathbb{P}_Y,$$

sean medidas sobre $(X, \mathcal{F}_X)$, para cada $A \in \mathcal{F}_X$. Así, podemos aplicar el teorema de convergencia monótona, ya que este teorema falla en el caso de las medidas con signo.

Definiremos la esperanza incondicional en **Prob** y probaremos que corresponde a una extensión natural de la definición de esperanza incondicional clásica.

Definición 2.30 (Esperanza Incondicional)**.** *Para $v \in \mathcal{L}(\overline{Y})$, llamaremos a $E^{!^-}(v)$ la esperanza incondicional de v, donde $!_{\overline{Y}}^- \colon \mathbb{0} \to \overline{Y}$ es la flecha en **Prob** definida en Proposición 2.7.*

Observación. Notemos que $E^{!_{\overline{Y}}^-}(v)$ corresponde a la esperanza condicional de v por $!_{\overline{Y}}^- \colon \mathbb{0} \to \overline{Y}$, es decir, $E^{!_{\overline{Y}}^-}(v) \in \{\{*\}, \emptyset\}$ es tal que

$$\int_{\{*\}} E^{!^-}(v) d\mathbb{P}_* = \int_{!_Y^{-1}(\{*\})} v\, d\mathbb{P}_Y.$$

La siguiente proposición establece que la esperanza incondicional en **Prob** corresponde a una generalización de la definición clásica.

Proposición 2.15. *Sean $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ un espacio de probabilidad, $v \in \mathcal{L}(\overline{Y})$ una variable aleatoria y $!_{\overline{Y}}^- \colon \mathbb{0} \to \overline{Y}$ la única flecha en **Prob**. Denotemos por $\mathbb{E}^{\mathbb{P}_Y}[v] = \int_Y v\, d\mathbb{P}_Y$, se tiene que*

$$E^{!_{\overline{Y}}^-}(v)(*) := \mathbb{E}^{\mathbb{P}_Y}[v].$$

Demostración. Sean $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ un espacio de probabilidad, $v \in \mathcal{L}(\overline{Y})$

una variable aleatoria y $!_Y^-\colon \mathbb{0} \to \overline{Y}$ la única flecha en **Prob**. Por la observación precedente y dado que $!_Y^{-1}(\{*\}) = Y$, tenemos que

$$\int_{\{*\}} E^{!^-}(v)\, d\mathbb{P}_* = \int_Y v\, d\mathbb{P}_Y. \tag{2.36}$$

Por otro lado, notemos que

$$\int_{\{*\}} E^{!^-}(v)\, d\mathbb{P}_* = E^{!^-}(v)(*) = E^{!^-}(v)(*),$$

de donde, junto con (2.36) se sigue que

$$E^{!^-}(v)(*) = \int_Y v\, d\mathbb{P}_Y = \mathbb{E}^{\mathbb{P}_Y}[v]. \qquad \square$$

Gracias al functor esperanza condicional podemos definir la medibilidad en la categoría **Prob**.

2.2.4. Medibilidad en Prob

En esta sección se define la medibilidad en **Prob**, la cual depende de una flecha f^- en **Prob**, además, se estudia la interacción entre la f^-−medibilidad y el functor esperanza condicional.

DEFINICIÓN 2.31 (f^--medibilidad). *Sean $f^-\colon \overline{X} \to \overline{Y}$ una flecha en **Prob** y $v \in \mathcal{L}^\infty(\overline{Y})$. Se dice que v es f^-−*medible *si existe $w \in \mathcal{L}^\infty(\overline{X})$ tal que*

$$v \sim_{\mathbb{P}_Y} w \circ f.$$

OBSERVACIÓN. Como podemos observar en la proposición 2.31 la noción de medibilidad en **Prob** depende de una flecha en **Prob**. Mientras que su contra parte, la medibilidad en teoría de probabilidades, depende de una σ−álgebra.

La siguiente proposición nos permite decir que un elemento de $L^1(\overline{Y})$ es f^-− medible.

PROPOSICIÓN 2.16. *Sean $f^-\colon \overline{X} \to \overline{Y}$ una flecha en **Prob** y $v_1, v_2 \in \mathcal{L}^\infty(\overline{Y})$ tales que $v_1 \sim_{\mathbb{P}_Y} v_2$. Si v_1 es f^-−medible, entonces v_2 también lo es.*

Demostración. Si v_1 es f^-–medible, sabemos existe $w_1 \in \mathcal{L}^\infty(\overline{X})$ tal que $v_1 \sim_{\mathbb{P}_Y} w \circ f$, lo cual es equivalente a tener que

$$v_1 = w_1 \circ f \quad \mathbb{P}_Y - \text{casi seguramente.} \tag{2.37}$$

Vamos a demostrar que existe $w_2 \in \mathcal{L}^\infty(\overline{X})$ tal que $v_2 \sim_{\mathbb{P}_Y} w_2 \circ f$. El resultado se sigue tomando $w_2 := w_1$. En efecto, dado que $v_1 = v_2$ $\mathbb{P}_Y$–casi seguramente, junto con (2.37), obtenemos que

$$v_2 = w_1 \circ f \quad \mathbb{P}_Y - \text{casi seguramente,}$$

lo cual es equivalente a tener que $v_2 \sim_{\mathbb{P}_Y} w_1 \circ f$. □

Para terminar esta sección estudiamos la interacción entre la f^-–medibilidad y la generalización de la esperanza condicional en **Prob**, mediante el siguiente teorema:

TEOREMA 2.17. *Sean $f^- \colon \overline{X} \to \overline{Y}$ una flecha en **Prob**, $u \in \mathcal{L}^1(\overline{Y})$ y $v \in \mathcal{L}^\infty(\overline{Y})$ una variable aleatoria. Supongamos que v es f^-–medible, se tiene que*

$$E^{f^-}(v \cdot u) \sim_{\mathbb{P}_X} w \cdot E^{f^-}(u),$$

donde $w \in \mathcal{L}^\infty(\overline{X})$ es una variable aleatoria que satisface $v \sim_{\mathbb{P}_Y} w \circ f$.

Demostración. Para probar el resultado, basta con demostrar que

$$\int_A E^{f^-}(v \cdot u)\, d\mathbb{P}_X = \int_A w E^{f^-}(u)\, d\mathbb{P}_X,$$

para cada $A \in \mathcal{F}_X$.

En primer lugar, consideremos la medida con signo

$$u^*(B) = \int_B u\, d\mathbb{P}_Y,$$

para cada $B \in \mathcal{F}_Y$, definida en $(Y, \mathcal{F}_Y)$. Así, tenemos que $u^* \circ f^{-1}$ es una medida con signo definida sobre $(X, \mathcal{F}_X)$.

La demostración la haremos en dos pasos, primero aplicaremos el teorema de descomposición de Jordan a la medida con signo u^*. Luego, aplicaremos el teorema de transformación A.6 a las medidas obtenidas en el primer paso.

1. Por el teorema de la descomposición de Jordan aplicado a la medida con signo u^* existen dos medidas mutuamente singulares, denotadas por $(u^*)^+$ y $(u^*)^-$, tales que

$$u^* = (u^*)^+ - (u^*)^- .$$

   Además, dado que $u = u^+ - u^-$, donde $u^+(x) = \max\{u(x), 0\}$ (parte positiva de u) y $u^-(x) = \min\{-u(x), 0\}$ (parte negativa de u), se tiene que

$$(u^*)^+ = \int_B u^+ \, d\mathbb{P}_Y \quad \text{y} \quad (u^*)^- = \int_B u^- \, d\mathbb{P}_Y,$$

   para cada $B \in \mathcal{F}_Y$.

   También, consideremos la medida con signo $u^* \circ f^{-1}$. Para cada $A \in \mathcal{F}_X$, tenemos que

   - $(u^* \circ f^{-1})^+ (A) = ((u^*)^+ \circ f^{-1}) (A) = \int_{f^{-1}(A)} u^+ \, d\mathbb{P}_X$;
   - $(u^* \circ f^{-1})^- (A) = ((u^*)^- \circ f^{-1}) (A) = \int_{f^{-1}(A)} u^- \, d\mathbb{P}_X$,

   las cuales son medidas definidas sobre $(X, \mathcal{F}_X)$ y satisfacen

$$u^* \circ f^{-1} = \left(u^* \circ f^{-1}\right)^+ - \left(u^* \circ f^{-1}\right)^- .$$

2. Para poder aplicar el teorema de transformación necesitamos que $(u^* \circ f^{-1})^+$ y $(u^* \circ f^{-1})^-$ sean absolutamente continuas con respecto a $\mathbb{P}_X$. En efecto, dado que $u \circ f^{-1} \ll \mathbb{P}_x$, por la proposición 1.2, se sigue que

$$\left(u^* \circ f^{-1}\right)^+ \ll \mathbb{P}_X \quad \text{y} \quad \left(u^* \circ f^{-1}\right)^- \ll \mathbb{P}_X.$$

   Sabemos que $w\colon (X, \mathcal{F}_X) \to \mathbb{R}, \mathbb{B}(\mathbb{R})$ es una función $\mathcal{F}_X/\mathcal{B}(\mathbb{R})-$medible y $f\colon (Y, \mathcal{F}_Y, (u^*)^+) \to (X, \mathcal{F}_X, \mathbb{P}_X)$ es una transformación $\mathcal{F}_Y/\mathcal{F}_X-$medible con $(X, \mathcal{F}_X, \mathbb{P}_X)$ un espacio medido finito. Así, por el teorema de transformación sabemos existe $\phi_1 = \frac{d((u^* \circ f^{-1})^+)}{d\mathbb{P}_X} = \left(\frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X}\right)^+ \in \mathcal{L}^1(\overline{X})$ tal que

$$\int_{f^{-1}(A)} w(f(y)) \, d(u^*)^+ = \int_A w(x)\phi_1(x) \, d\mathbb{P}_X,$$

para cada $A \in \mathcal{F}_X$, de donde, por la definición de la medida $(u^*)^+$, se sigue que

$$\int_{f^{-1}(A)} w(f(y))u^+ \, d\mathbb{P}_Y = \int_A w(x)\phi_1(x) \, d\mathbb{P}_X. \tag{2.38}$$

Análogamente para la medida $(u^*)^-$, por el teorema de transformación, sabemos existe $\phi_2 = \frac{d((u^* \circ f^{-1})^-)}{d\mathbb{P}_X} = \left(\frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X}\right)^- \in \mathcal{L}^1(\overline{X})$ tal que

$$\int_{f^{-1}(A)} w(f(y))u^- \, d\mathbb{P}_Y = \int_A w(x)\phi_2(x) \, d\mathbb{P}_X, \tag{2.39}$$

para cada $F \in \mathcal{F}_X$.

Por otro lado, como v es f^- medible existe $w \in \mathcal{L}^\infty(X)$ tal que $v \sim_{\mathbb{P}_Y} w \circ f$. Así, la esperanza condicional de $v \cdot u$ por f^- satisface que

$$\begin{aligned}
\int_A E^{f^-}(v \cdot u) \, d\mathbb{P}_X &= \int_{f^{-1}(A)} v \cdot u \, d\mathbb{P}_Y, \\
&= \int_{f^{-1}(A)} (w \circ f) \cdot u \, d\mathbb{P}_Y, \\
&= \int_{f^{-1}(A)} (w \circ f) \cdot u^+ \, d\mathbb{P}_Y - \int_{f^{-1}(A)} (w \circ f) \cdot u^- \, d\mathbb{P}_Y,
\end{aligned}$$

para $A \in \mathcal{F}_X$, de donde, junto con (2.38) y (2.39), se sigue que

$$\begin{aligned}
\int_A E^{f^-}(v \cdot u) \, d\mathbb{P}_X &= \int_A w \cdot \phi_1 \, d\mathbb{P}_X - \int_A w \cdot \phi_2 \, d\mathbb{P}_X, \\
&= \int_A w \cdot (\phi_1 - \phi_2) \; d\mathbb{P}_X, \\
&= \int_A w \cdot \left(\left(\frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X} \right)^+ - \left(\frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X} \right)^- \right) \, d\mathbb{P}_X,
\end{aligned}$$

para $A \in \mathcal{F}_X$, luego, dado que $E^{f^-}(u) = \frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X} = \left(\frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X}\right)^+ - \left(\frac{d(u^* \circ f^{-1})}{d\mathbb{P}_X}\right)^-$, obtenemos que

$$\int_A E^{f^-}(v \cdot u) \, d\mathbb{P}_X = \int_A w E^{f^-}(u) \, d\mathbb{P}_X,$$

para $A \in \mathcal{F}_X$, con lo cual, podemos concluir que

$$E^{f^-}(v \cdot u) \sim_{\mathbb{P}_X} w \cdot E^{f^-}(u),$$

como queríamos. $\square$

2.2.5. Independencia de flechas en Prob

En esta sección se extiende la noción de independencia en la categoría **mpProb** a la categoría **Prob**.

Además, consideraremos a $\overline{X} \otimes \overline{Y} = (X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y, \mathbb{P}_X \otimes \mathbb{P}_Y)$ como el espacio de probabilidad producto entre $\overline{X}$ y $\overline{Y}$, el cual se definió en la sección A.3 del Capítulo 1.

Antes de extender esta noción de independencia a la categoría **Prob** es necesario considerar la siguiente proposición:

PROPOSICIÓN 2.18. *Sea $f^-: \overline{X} \to \overline{Y}$ una flecha en **Prob**. Consideremos el objeto $\overline{X}_{f^-}$ en **Prob** definido por $\overline{X}_{f^-} = (X, \mathcal{F}_X, \mathbb{P}_Y \circ f^{-1})$. Se tiene que el siguiente diagrama:*

$$\begin{tikzcd} \overline{X} \arrow[r, "f^-"] \arrow[d, "id^-_{\overline{X}}"'] & \overline{Y} \\ \overline{X}_f \arrow[ur, "f^\sim"'] & \end{tikzcd}$$

*conmuta en **Prob**, es decir,*

$$f^\sim \circ id^-_{\overline{X}} = f^-,$$

donde $id^-_{\overline{X}}$ y $f^\sim$ son las flechas correspondientes a las funciones $id_X: \overline{X}_{f^-} \to \overline{X}$ y $f: \overline{Y} \to \overline{X}_{f^-}$, respectivamente. Además, se tiene que $f^\sim$ preserva medidas.

Demostración. Por definición de composición en **Prob**, tenemos que

$$f^\sim \circ id^-_{\overline{X}} = (id_X \circ f)^- = f^-.$$

Notemos que $f^{\sim}\colon \overline{X}_f \to \overline{Y}$ preserva medidas, pues se cumple que

$$\mathbb{P}_{\overline{X}_f} = \mathbb{P}_Y \circ f^{-1}. \qquad \square$$

Con esto, presentamos la extensión de independencia en la categoría **Prob**:

Definición 2.32 (Independencia en **Prob**). *Sean $f^{-}\colon \overline{X} \to \overline{Z}$ y $g^{-}\colon \overline{Y} \to \overline{Z}$ dos flechas en **Prob**. Se dice que f^{-} y g^{-} son independientes si existe $q^{-}\colon \overline{X}_{f^{-}} \otimes \overline{Y}_{g^{-}} \to \overline{Z}$, una flecha que preserva medidas, tal que el siguiente diagrama:*

$$\begin{tikzcd}
\overline{X} \arrow[r, "f^{-}"] \arrow[d, "id_X^{-}"'] & \overline{Z} & \overline{Y} \arrow[l, "g^{-}"'] \arrow[d, "id_Y^{-}"] \\
\overline{X}_f \arrow[r, "p_1^{-}"'] \arrow[ur, "f^{\sim}"] & \overline{X}_{f^{-}} \otimes \overline{Y}_{g^{-}} \arrow[u, dashed, "q^{-}"'] & \overline{Y}_g \arrow[l, "p_2^{-}"] \arrow[ul, "g^{\sim}"']
\end{tikzcd}$$

conmuta, es decir,

$$q^{-} \circ p_1^{-} = f^{\sim} \quad \text{y} \quad q^{-} \circ p_2^{-} = g^{\sim}. \tag{2.40}$$

Observación. Notemos que 2.40 de la definición 2.32 puede reescribirse como

$$q^{-} \circ p_1^{-} = f^{-} \quad \text{y} \quad q^{-} \circ p_2^{-} = g^{-}.$$

En efecto, por un lado, notemos que

$$\begin{aligned}
(q^{-} \circ p_1^{-}) \circ id_X^{-} &= (p_1 \circ q)^{-} \circ id_X^{-}, \\
&= ((id_X \circ p_1) \circ q)^{-}, \\
&= (p_1 \circ q)^{-}, \\
&= q^{-} \circ p_1^{-}.
\end{aligned}$$

Por otro lado, de la proposición 2.18, tenemos que

$$f^{\sim} \circ id_X^{-} = f^{-},$$

con lo cual, podemos concluir que

$$q^{-} \circ p_1^{-} = f^{-}.$$

Análogamente, se tiene que

$$q^{-} \circ p_2^{-} = g^{-}.$$

Observación. Para presentar los resultados posteriores consideraremos la inyección canónica de $L^{\infty}(\overline{X})$ a $L^{1}(\overline{X})$, definida por

$$\iota : L^{\infty}(\overline{X}) \hookrightarrow L^{1}(\overline{X}), \quad \iota([u]_{\sim\mathbb{P}_{\mathbb{X}}}) = [u]_{\sim\mathbb{P}_{\mathbb{X}}},$$

donde $[u]_{\sim\mathbb{P}_X} \in L^{\infty}(\overline{X})$.

Previo a estudiar la interacción entre la independencia y el functor esperanza condicional consideremos los siguientes lemas:

Lema 2.19. *Sea $f^{-}: \overline{X} \to \overline{Y}$ una flecha en **Prob** que preserva medidas, se tiene que*

$$\boldsymbol{\varepsilon}(f^{-}) \circ \iota_1 \circ \mathbf{L}(f^{-}) = \iota,$$

donde $\iota : L^{\infty}(\overline{X}) \hookrightarrow L^{1}(\overline{X})$ y $\iota_1 : L^{\infty}(\overline{Y}) \hookrightarrow L^{1}(\overline{Y})$.

Demostración. Sea $u \in \mathcal{L}^{\infty}(\overline{X})$, vamos a demostrar que

$$\big(\boldsymbol{\varepsilon}(f^{-}) \circ \mathbf{L}(f^{-})\big)\left([u]_{\sim\mathbb{P}_X}\right) = \iota\left([u]_{\sim\mathbb{P}_X}\right).$$

En primer lugar, aplicando la definición del functor $\mathbf{L}$ para flechas, se tiene que

$$\begin{aligned} \mathbf{L}(f^{-}) :& \mathrm{L}^{\infty}(\overline{X}) \longrightarrow \mathrm{L}^{\infty}(\overline{Y}) \\ & [u]_{\sim\mathbb{P}_X} \longmapsto \left(\mathbf{L}(f^{-})\right)\left([u]_{\sim\mathbb{P}_X}\right) = [u \circ f]_{\sim\mathbb{P}_Y}, \end{aligned}$$

luego, consideremos la inyección canónica $\iota_1 : \mathrm{L}^{\infty}(\overline{Y}) \hookrightarrow \mathrm{L}^{1}(\overline{Y})$, de este modo, tenemos que

$$\begin{aligned} \iota_1 \circ \mathbf{L}(f^{-}) :& \mathrm{L}^{\infty}(\overline{X}) \longrightarrow \mathrm{L}^{1}(\overline{Y}) \\ & [u]_{\sim\mathbb{P}_X} \longmapsto \iota_1\left([u \circ f]_{\sim\mathbb{P}_Y}\right) = [u \circ f]_{\sim\mathbb{P}_Y}. \end{aligned} \tag{2.41}$$

Por otro lado, sabemos que

$$\begin{aligned} \boldsymbol{\varepsilon}(f^-) \colon \mathrm{L}^\infty(\overline{Y}) &\longrightarrow \mathrm{L}^1(\overline{X}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto [E^{f^-}(u)]_{\sim\mathbb{P}_X}. \end{aligned}$$

En particular, la esperanza condicional de $u \circ f$ por f^-, denotada por $E^{f^-}(u \circ f)$ es tal que

$$\int_A E^{f^-}(u \circ f)\, d\mathbb{P}_X = \int_{f^{-1}(A)} u \circ f\, d\mathbb{P}_Y,$$

para cada $A \in \mathcal{F}_X$, luego, por el teorema A.4, se tiene que

$$\int_A E^{f^-}(u \circ f)\, d\mathbb{P}_X = \int_{f^{-1}(A)} u \circ f\, d\mathbb{P}_Y = \int_A u\, d(\mathbb{P}_Y \circ f^{-1}),$$

para cada $A \in \mathcal{F}_X$, con esto, dado que $\mathbb{P}_X = \mathbb{P}_Y \circ f^{-1}$, obtenemos que

$$\int_A E^{f^-}(u \circ f)\, d\mathbb{P}_X = \int_A u\, d\mathbb{P}_X,$$

para cada $A \in \mathcal{F}_X$. Así, podemos concluir que

$$\boldsymbol{\varepsilon}(f^-)\left([u \circ f]_{\sim\mathbb{P}_X}\right) = [E^{f^-}(u \circ f)]_{\sim\mathbb{P}_X} = [u]_{\sim\mathbb{P}_X},$$

con lo cual, junto con (2.41), se sigue que

$$\begin{aligned} \boldsymbol{\varepsilon}(f^-) \circ \iota_1 \circ \mathbf{L} \colon \mathrm{L}^\infty(\overline{X}) &\longrightarrow \mathrm{L}^1(\overline{X}) \\ [u]_{\sim\mathbb{P}_X} &\longmapsto [u]_{\mathbb{P}_X}. \end{aligned}$$

Es decir, hemos probado que

$$\boldsymbol{\varepsilon}(f^-) \circ \iota_1 \circ \mathbf{L}(f^-) = \iota.$$

como queríamos. □

Lema 2.20. *Sean* $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ *y* $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ *espacios de probabilidad. Para* $v \in \mathcal{L}^\infty(\overline{X})$*, se tiene que*

$$E^{p_1^-}(v \circ p_2) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[v] E^{Id_X^-}(\mathbb{1}_X),$$

donde $p_1 \colon X \times Y \to X$ *y* $p_2 \colon X \times Y \to Y$ *son las funciones proyecciones.*

Demostración. La esperanza condicional de $v \circ p_2 \colon X \times Y \to \mathbb{R}$ por $p_1^- : \overline{X} \to \overline{X} \otimes \overline{Y}$ satisface

$$\int_A E^{f^-}(v \circ p_2)\, d\mathbb{P}_X = \int_{p_1^{-1}(A)} (v \circ p_2)\, d(\mathbb{P}_X \otimes \mathbb{P}_Y),$$

para cada $A \in \mathcal{F}_X$, de donde, dado que $p^{-1}(A) = A \times Y$, tenemos que

$$\begin{aligned}
\int_A E^{f^-}(v \circ p_2)\, d\mathbb{P}_X &= \int_{A\times Y} (v \circ p_2)\, d(\mathbb{P}_X \otimes \mathbb{P}_Y),\\
&= \int_{X\times Y} (v \circ p_2) \cdot \mathbb{1}_{A\times Y}\, d(\mathbb{P}_X \otimes \mathbb{P}_Y),\\
&= \int_{X\times Y} (v \circ p_2) \cdot (\mathbb{1}_A \circ p_1)\, d(\mathbb{P}_X \otimes \mathbb{P}_Y),
\end{aligned}$$

para cada $A \in \mathcal{F}_X$, de donde, aplicando el teorema de Fubini A.5, obtenemos que

$$\begin{aligned}
\int_A E^{f^-}(v \circ p_2)\, d\mathbb{P}_X &= \int_Y \int_X ((v \circ p_2)(x,y)) \cdot ((\mathbb{1}_A \circ p_1(x,y)))\, \mathbb{P}_X(dx)\, \mathbb{P}_Y(dy),\\
&= \int_Y \int_X v(y) \cdot \mathbb{1}_A(x)\, \mathbb{P}_X(dx)\, \mathbb{P}_Y(dy),\\
&= \int_Y v(y)\, \mathbb{P}_Y(dy) \int_X \mathbb{1}_A(x)\, \mathbb{P}_X(dx),\\
&= \mathbb{E}^{\mathbb{P}_Y}[v] \int_A \mathbb{1}_X\, d\mathbb{P}_X,\\
&= \mathbb{E}^{\mathbb{P}_Y}[v] \int_A E^{Id_X^-}(\mathbb{1}_X)\, d\mathbb{P}_X,
\end{aligned}$$

para cada $A \in \mathcal{F}_X$. Así, podemos concluir que

$$E^{p_1^-}(v \circ p_2) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[v] E^{Id_X^-}(\mathbb{1}_X),$$

como queríamos. □

Observación. El resultado del lema 2.20 puede reescribirse como

$$E^{p_1^-}(v \circ p_2) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[v] \mathbb{1}_X.$$

En efecto, para el caso particular de la flecha identidad en **Prob** se tiene que

$$\int_A \mathbb{1}_X\, d\mathbb{P}_X = \int_A E^{Id_X^-}(\mathbb{1}_X)\, d\mathbb{P}_X,$$

para cada $A \in \mathcal{F}_X$, con lo cual podemos concluir que

$$\mathbb{1}_X \sim_{\mathbb{P}_X} E^{Id_X^-}(\mathbb{1}_X).$$

Con esto, presentamos el siguiente teorema que establece la relación entre el funtor esperanza condicional y la esperanza condicional usual:

TEOREMA 2.21. *Sean $f^- : \overline{X} \to \overline{Y}$ y $g^- : \overline{Y} \to \overline{Y}$ dos flechas independientes en **Prob**. Para $v \in \mathcal{L}^\infty(\overline{X})$, se tiene que*

$$E^{f^-}(v \circ g) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Z}(v \circ g) E^{f^-}(\mathbb{1}_Z).$$

Demostración. La demostración la haremos en dos pasos, primero aplicaremos los functores ε y $\mathbf{L}$ al diagrama de independencia de las flechas f^- y g^-. Luego, probaremos que el diagrama resultante conmuta. Finalmente, tomaremos dos caminos de este diagrama y colegiremos el resultado.

1. Dado que $f^- : \overline{X} \to \overline{Y}$ y $g^- : \overline{Y} \to \overline{Y}$ son flechas independientes en **Prob** existe $q^- : \overline{X}_f \otimes \overline{Y}_f \to \overline{Z}$ una flecha que preserva medidas tal que el siguiente diagrama:

$$\begin{tikzcd}
\overline{X} \arrow[r, "f^-"] \arrow[d, "id_X^-"'] & \overline{Z} & \overline{Y} \arrow[l, "g^-"'] \arrow[d, "id_Y^-"] \\
\overline{X}_{f^-} \arrow[r, "p_1^-"'] \arrow[ur, "f^\sim"] & \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \arrow[u, dashed, "q^-"'] & \overline{Y}_{g^-} \arrow[l, "p_2^-"] \arrow[ul, "g^\sim"']
\end{tikzcd}$$

   conmuta.

   Ahora, apliquemos el functor ε al cuadrado izquierdo del diagrama y el functor $\mathbf{L}$ al cuadrado derecho del diagrama, con esto, tenemos que

   Vamos a relacionar estos diagramas empleando las inyecciones canónicas:

$$\iota_1 : L^\infty(\overline{Z}) \hookrightarrow L^1(\overline{Z}) \quad \text{y} \quad \iota_2 : L^\infty(\overline{X}_{f^-} \otimes \overline{Y}_{g^-}) \hookrightarrow L^1(\overline{X}_{f^-} \otimes \overline{Y}_{g^-}).$$

$$
\begin{tikzcd}
\mathrm{L}^1(\overline{X}) \arrow[d, "\boldsymbol{\varepsilon}(id_X^-)"'] & \mathrm{L}^1(\overline{Z}) \arrow[l, "\boldsymbol{\varepsilon}(f^-)"'] \arrow[dl, "\boldsymbol{\varepsilon}(f^\sim)"'] \arrow[d, dashed, "\boldsymbol{\varepsilon}(q^-)"] \\
\mathrm{L}^1(\overline{X}_{f^-}) \arrow[r, "\boldsymbol{\varepsilon}(p_1^-)"'] & \mathrm{L}^1(\overline{X}_{f^-} \otimes \overline{Y}_{g^-})
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{L}^\infty(\overline{Z}) & \mathrm{L}^\infty(\overline{Y}) \arrow[l, "\mathbf{L}(g^-)"'] \arrow[d, "\mathbf{L}(id_Y^-)"] \\
\mathrm{L}^\infty(\overline{X}_{f^-} \otimes \overline{Y}_{g^-}) \arrow[u, dashed, "\mathbf{L}(q^-)"] & \mathrm{L}^\infty(\overline{Y}_{g^-}) \arrow[l, "\mathbf{L}(p_2^-)"] \arrow[ul, "\mathbf{L}(g^\sim)"']
\end{tikzcd}
$$

En efecto, tenemos que

$$
\begin{tikzcd}
\mathrm{L}^1(\overline{X}) & \mathrm{L}^1(\overline{Z}) \arrow[l, "\boldsymbol{\varepsilon}(f^-)"'] \arrow[dl, "\boldsymbol{\varepsilon}(f^\sim)"'] \arrow[d, dashed, "\boldsymbol{\varepsilon}(q^-)"] & \mathrm{L}^\infty(\overline{Z}) \arrow[l, "\iota_1"'] & \mathrm{L}^\infty(\overline{Y}) \arrow[l, "\mathbf{L}(g^-)"'] \arrow[d, "\mathbf{L}(id_Y^-)"] \\
\mathrm{L}^1(\overline{X}_{f^-}) \arrow[u, "\boldsymbol{\varepsilon}(id_X^-)"] & \mathrm{L}^1(\overline{X}_{f^-} \otimes \overline{Y}_{g^-}) \arrow[l, "\boldsymbol{\varepsilon}(p_1^-)"] & \mathrm{L}^\infty(\overline{X}_{f^-} \otimes \overline{Y}_{g^-}) \arrow[l, "\iota_2"] \arrow[u, dashed, "\mathbf{L}(q^-)"] & \mathrm{L}^\infty(\overline{Y}_{g^-}) \arrow[l, "\mathbf{L}(p_2^-)"] \arrow[ul, "\mathbf{L}(g^\sim)"']
\end{tikzcd}
$$

Figura 2.1: Diagrama final.

Por otro lado, sabemos que

$$
q^- \circ p_1^- = f^\sim \quad \text{y} \quad q^- \circ p_2^- = g^\sim,
$$

de donde, puesto que ε y $\mathbf{L}$ preservan la composición, se tiene que

$$
\boldsymbol{\varepsilon}(p_1^-) \circ \boldsymbol{\varepsilon}(q^-) = \boldsymbol{\varepsilon}(f^\sim) \quad \text{y} \quad \mathbf{L}(q^-) \circ \mathbf{L}(p_2^-) = \mathbf{L}(g^\sim), \tag{2.42}
$$

es decir, de 2.42 tenemos que los cuadrados izquierdo y derecho en el diagrama anterior conmutan. Gracias al Lema 2.19, el cuadrado central también conmuta. En efecto, sea $[u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}} \in \mathrm{L}^\infty\left(\overline{X}_{f^-} \otimes \overline{Y}_{g^-}\right)$, vamos a demostrar que

$$
\left(\boldsymbol{\varepsilon}(q^-) \circ \iota_1 \circ \mathbf{L}(q^-)\right)\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right) = \iota_2\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right).
$$

Por un lado, notemos que

$$\begin{aligned}\left(\boldsymbol{\varepsilon}(q^-) \circ \iota_1 \circ \mathbf{L}(q^-)\right)\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right) &= \left(\boldsymbol{\varepsilon}(q^-)\right) \circ \iota_1 \left(\mathbf{L}(q^-)\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right)\right) \\ &= \left(\boldsymbol{\varepsilon}(q^-)\right) \circ \left(\mathbf{L}(q^-)\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right)\right),\end{aligned}$$

de donde, junto con el Lema 2.19, se sigue que

$$\left(\boldsymbol{\varepsilon}(q^-) \circ \iota_1 \circ \mathbf{L}(q^-)\right)\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right) = id_{\mathrm{L}^\infty(\overline{X}_{f^-} \otimes \overline{Y}_{g^-})}\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right) = [u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}.$$

Por otro lado, tenemos que

$$\iota_2\left([u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}\right) = [u]_{\sim_{\mathbb{P}_X \otimes \mathbb{P}_Y}}.$$

Así, podemos concluir que

$$\boldsymbol{\varepsilon}(q^-) \circ \iota_1 \circ \mathbf{L}(q^-) = \iota_2,$$

como queríamos.

2. En resumen, hemos probado que el diagrama 2.1 conmuta. En particular, para $v \in \mathcal{L}^\infty(\overline{Y})$, analicemos los valores en la esquina superior izquierda del diagrama 2.1 desarrollados a través de dos caminos, que deberían coincidir. El primer camino viene dado por:

$$\begin{aligned}\left(\boldsymbol{\varepsilon}(f^-) \circ \iota_1 \circ \mathbf{L}(g^-)\right)\left([v]_{\sim \mathbb{P}_Y}\right) &= \left(\boldsymbol{\varepsilon}(f^-) \circ \iota_1\right)\left([v \circ g]_{\sim \mathbb{P}_Y}\right) \\ &= \boldsymbol{\varepsilon}(f^-)\left([v \circ g]_{\sim \mathbb{P}_Y}\right) \\ &= [E^f(v \circ g)]_{\sim \mathbb{P}_X}.\end{aligned}$$

Como notación consideremos $\psi := \left(\boldsymbol{\varepsilon}(id_X^-) \circ \boldsymbol{\varepsilon}(p_1^-) \circ \iota_2 \circ \mathbf{L}(p_2^-) \circ \mathbf{L}(id_Y^-)\right)$. Así, el segundo camino es:

$$\begin{aligned}\psi\left([v]_{\sim \mathbb{P}_Y}\right) &= \left(\boldsymbol{\varepsilon}(id_X^-) \circ \boldsymbol{\varepsilon}(p_1^-) \circ \iota_2 \circ \mathbf{L}(p_2^-)\right)\left([v]_{\sim \mathbb{P}_Y}\right) \\ &= \left(\boldsymbol{\varepsilon}(id_X^-) \circ \boldsymbol{\varepsilon}(p_1^-) \circ \iota_2\right)\left([v \circ p_2]_{\sim \mathbb{P}_{\overline{X}_f^-} \otimes \mathbb{P}_{\overline{Y}_{g^-}}}\right) \\ &= \left(\boldsymbol{\varepsilon}(id_X^-) \circ \boldsymbol{\varepsilon}(p_1^-)\right)\left([v \circ p_2]_{\sim \mathbb{P}_{\overline{X}_{f^-} \otimes \overline{Y}_{g^-}}}\right) \\ &= \boldsymbol{\varepsilon}(id_X^-)\left([E^{p_1^-}(v \circ p_2)]_{\mathbb{P}_{\overline{X}_f^-}}\right)\end{aligned}$$

$$= \left[E^{id_X^-} \left(E^{p_1^-}(v \circ p_2) \right) \right]_{\sim \mathbb{P}_X},$$

de donde, gracias a la proposición 2.10, se sigue que

$$\left(\boldsymbol{\varepsilon}(id_X^-) \circ \boldsymbol{\varepsilon}(p_1^-) \circ \iota_2 \circ \mathbf{L}(p_2^-) \circ \mathbf{L}(id_Y^-) \right)([v]_{\sim \mathbb{P}_Y}) = [E^{p_1^-}(v \circ p_2)]_{\sim \mathbb{P}_X}.$$

Con esto, tenemos que

$$[E^f(v \circ g)]_{\sim \mathbb{P}_X} = [E^{p_1^-}(v \circ p_2)]_{\sim \mathbb{P}_X}. \tag{2.43}$$

Recordemos que $p_1^- : \overline{X}_f^- \to \overline{X}_f^- \otimes \overline{Y}_g^-$ y $p_2^- : \overline{X}_f^- \to \overline{Y}_f^- \otimes \overline{Y}_g^-$, donde $\overline{X}_f^- = (X, \mathcal{F}_X, \mathbb{P}_Z \circ f^{-1})$ y $\overline{Y}_f^- = (Y, \mathcal{F}_Y, \mathbb{P}_Z \circ g^{-1})$. Así, por el lema 2.20, obtenemos que

$$E^{p_1^-}(v \circ p_2) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Z \circ g^{-1}}(v) E^{id_X^-}(\mathbb{1}_X). \tag{2.44}$$

Ahora, para cada $A \in \mathcal{F}_X$, notemos que

$$\begin{aligned}
\int_A E^{id_X^-}(\mathbb{1}_X)\, d\mathbb{P}_X &= \int_{id_X^{-1}(A)} \mathbb{1}_X \, d\mathbb{P}_Z \circ f^{-1}, \\
&= \int_A \mathbb{1}_X \, d\mathbb{P}_Z \circ f^{-1}, \\
&= \left(\mathbb{P}_Z \circ f^{-1} \right)(A), \\
&= \mathbb{P}_Z \left(f^{-1}(A) \right), \\
&= \int_{f^{-1}(A)} \mathbb{1}_Z \, d\mathbb{P}_Z, \\
&= \int_A E^{f^-}[\mathbb{1}_Z] \, d\mathbb{P}_X,
\end{aligned}$$

de donde, podemos concluir que

$$E^{id_X^-}(\mathbb{1}_X) \sim_{\mathbb{P}_X} E^{f^-}(\mathbb{1}_Z).$$

También, por la fórmula de cambio de variable, se tiene que

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}_Z \circ g^{-1}}(v) &= \int_Y v \, d\left(\mathbb{P}_Z \circ g^{-1} \right), \\
&= \int_{g^-(Y)} v \circ g \mathbb{P}_Z, \\
&= \mathbb{E}^{\mathbb{P}_Z}[v \circ g].
\end{aligned}$$

Así, la ecuación (2.44) puede reescribirse como sigue

$$E^{p_1^-}(v \circ p_2) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Z}[v \circ g]E^{f^-}(\mathbb{1}_Z),$$

luego, junto con (2.43), se sigue que

$$E^f(v \circ g) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Z}[v \circ g]E^{f^-}(\mathbb{1}_Z),$$

como queríamos. $\square$

Finalmente, vamos a presentar la definición de independencia con respecto a una flecha en **Prob**.

Definición 2.33 (f^-–independencia). *Para una variable aleatoria $v \in \mathcal{L}^1(\overline{Y})$, definimos el espacio de probabilidad $\overline{\mathbb{R}}_v$ por $\overline{\mathbb{R}}_v = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_Y \circ v^{-1})$. Se tiene que $v^- \colon \overline{R}_v \to \overline{Y}$ es una flecha en **Prob**. Ahora, para una flecha en **Prob** $f^- \colon \overline{X} \to \overline{Y}$, se dice que v es independiente a f^-, denotado por $v \perp f^-$, si f^- y v^- son independientes.*

Observación. Sabemos que $v \colon \overline{Y} \to \overline{R}_v$ es $\mathcal{F}_Y/\mathcal{B}(\mathbb{R})$–medible, basta con demostrar que v es nula-conservadora, para lo cual, debemos probar que

$$\mathbb{P}_Y \circ v^{-1} \ll \mathbb{P}_{\overline{R}_v}.$$

El resultado se sigue dado que por definición $\mathbb{P}_{\overline{R}_v} = \mathbb{P}_Y \circ v^{-1}$.

La siguiente proposición nos permite decir que un elemento de $\mathrm{L}^1(\overline{Y})$ es independiente a f^-.

Proposición 2.22. *Sean $f^- \colon \overline{X} \to \overline{Y}$ una flecha en **Prob** y $v_1, v_2 \in \mathcal{L}^1(\overline{Y})$ tales que $v_1 \sim_{\mathbb{P}_X} v_2$. Se tiene que $v_1 \perp f$ implica que $v_2 \perp f$.*

Demostración. En primer lugar, notemos que $v_1^- \colon \overline{\mathbb{R}}_{v_1} \to \overline{Y}$ y $v_2^- \colon \overline{\mathbb{R}}_{v_2} \to \overline{Y}$ corresponden a las flechas en **Prob** de $v_1 \colon \overline{Y} \to \overline{\mathbb{R}}_{v_2}$ y $v_2 \colon \overline{Y} \to \overline{\mathbb{R}}_{v_2}$, respectivamente.

Supongamos que $v_1 \perp f$, es decir, f^- y v^- son independientes. Así, sabemos existe $q_1 \colon \overline{X}_f^- \otimes \overline{\mathbb{R}}_{v_1} \to \overline{Y}$ una flecha en **Prob** que preserva medidas tal que el siguiente diagrama:

$$
\begin{tikzcd}
\overline{X} \arrow[r, "f^{-}"] \arrow[d, "id_X^{-}"'] & \overline{Y} & \\
\overline{X}_{f^{-}} \arrow[r, "p_1^{-}"'] \arrow[ur, "f^{\sim}"] & \overline{X}_{f^{-}} \otimes \overline{\mathbb{R}}_{v_1} \arrow[u, dashed, "q_1^{-}"'] & \overline{\mathbb{R}}_{v_1} \arrow[l, "p_2^{-}"] \arrow[ul, "v_1^{-}"']
\end{tikzcd}
$$

conmuta. Con esto, se tiene que

$$
\begin{aligned}
q_1 \colon \overline{Y} &\longrightarrow \overline{X}_{f^{-}} \otimes \overline{\mathbb{R}}_{v_1} \\
y &\longmapsto q(y) = (f(y), v_1(y)).
\end{aligned}
$$

Vamos a demostrar que existe $q_2 \colon \overline{X}_{f^{-}} \otimes \overline{\mathbb{R}}_{v_2} \to \overline{Y}$ una flecha en **Prob** que preserva medidas tal que el siguiente diagrama:

$$
\begin{tikzcd}
\overline{X} \arrow[r, "f^{-}"] \arrow[d, "id_X^{-}"'] & \overline{Y} & \\
\overline{X}_{f^{-}} \arrow[r, "p_1^{-}"'] \arrow[ur, "f^{\sim}"] & \overline{X}_{f^{-}} \otimes \overline{\mathbb{R}}_{v_2} \arrow[u, dashed, "q_2^{-}"'] & \overline{\mathbb{R}}_{v_2} \arrow[l, "p_2^{-}"] \arrow[ul, "v_2^{-}"']
\end{tikzcd}
$$

conmuta.

Para ello, primero probemos que $\overline{\mathbb{R}}_{v_1} = \overline{\mathbb{R}}_{v_2}$. Notemos que es suficiente demostrar que

$$
\mathbb{P}_Y \circ v_1^{-1} = \mathbb{P}_Y \circ v_2^{-1}.
$$

Sea $B \in \mathcal{B}(\mathbb{R})$, tenemos que

$$
\left(\mathbb{P}_Y \circ v_1^{-1}\right)(B) = \mathbb{P}_Y(v_1^{-1}(B)). \tag{2.45}
$$

Por otro lado, por hipótesis se tiene que $v_1 \sim_{\mathbb{P}_X} v_2$, con lo cual podemos considerar el siguiente conjunto:

$$
N = \{y \in Y : v_1(y) \neq v_2(y)\},
$$

el cual satisface que $\mathbb{P}_Y(N) = 0$. Además, denotemos por $M := Y \setminus N$. Así, obtenemos que

$$v_1^{-1}(B) = \left(v_1^{-1}(B) \cap N\right) \cup \left(v_1^{-1}(B) \cap M\right),$$

de donde, de (2.45) y por la aditividad de $\mathbb{P}_Y$, se sigue que

$$\begin{aligned}\left(\mathbb{P}_Y \circ v_1^{-1}\right)(B) &= \mathbb{P}_Y(v_1^{-1}(B)), \\ &= \mathbb{P}_Y\left(\left(v_1^{-1}(B) \cap N\right) \cup \left(v_1^{-1}(B) \cap M\right)\right), \\ &= \mathbb{P}_Y\left(\left(v_1^{-1}(B) \cap N\right)\right) + \mathbb{P}_Y\left(\left(v_1^{-1}(B) \cap M\right)\right).\end{aligned}$$

Luego, dado que $v_1^{-1}(B) \cap N \subset N$, se tiene que

$$0 \leq \mathbb{P}_Y(v_1^{-1}(B)) \leq \mathbb{P}_Y(N) = 0,$$

es decir, $\mathbb{P}_Y(v_1^{-1}(B)) = 0$. Por lo tanto, tenemos que

$$\left(\mathbb{P}_Y \circ v_1^{-1}\right)(B) = \mathbb{P}_Y\left(\left(v_1^{-1}(B) \cap M\right)\right). \tag{2.46}$$

De igual forma, se prueba que

$$\left(\mathbb{P}_Y \circ v_2^{-1}\right)(B) = \mathbb{P}_Y\left(\left(v_2^{-1}(B) \cap M\right)\right). \tag{2.47}$$

Ahora, sea $y \in v_1^{-1}(B) \cap M$, tenemos las siguientes equivalencias:

$$y \in v_1^{-1}(B) \cap M \iff v_1(y) = v_2(y) \in B \iff y \in v_2^{-1}(B) \cap M,$$

así, hemos probado que

$$v_1^{-1}(B) \cap M = v_2^{-1}(B) \cap M.$$

de donde, junto con (2.46) y (2.47), se sigue que

$$\left(\mathbb{P}_Y \circ v_1^{-1}\right)(B) = \left(\mathbb{P}_Y \circ v_2^{-1}\right)(B).$$

Con esto, podemos concluir que $\overline{\mathbb{R}}_{v_1} = \overline{\mathbb{R}}_{v_2}$. Así, basta definir q_2^- por

$$\begin{aligned} q_2 : \overline{Y} &\longrightarrow \overline{X}_{f^-} \otimes \overline{\mathbb{R}}_{v_2} \\ y &\longmapsto q(y) = (f(y), v_2(y)),\end{aligned}$$

de donde, se cumple que el diagrama de independencia de v_2 y f conmuta.

Finalmente, probemos que q_2^- preserva medidas, para lo cual, debemos demostrar que

$$\mathbb{P}_{\overline{X}_{f^-}} \otimes \mathbb{P}_{\overline{\mathbb{R}}_{v_2}} = \mathbb{P}_Y \circ f^{-1} \otimes \mathbb{P}_Y \circ v_2^{-1} = \mathbb{P}_Y \circ q_2^{-1},$$

lo cual, dado que $\mathbb{P}_Y \circ v_1^{-1} = \mathbb{P}_Y \circ v_2^{-1}$, es equivalente a probar que

$$\mathbb{P}_Y \circ q_1^{-1} = \mathbb{P}_Y \circ q_2^{-1}.$$

Sea $E \in \mathcal{F}_X \otimes \mathcal{B}(\mathbb{R})$, tenemos que

$$\begin{aligned}
\mathbb{P}_Y \circ q_1^{-1}(E) &= \mathbb{P}_Y \left(q_1^{-1}(E)\right), \\
&= \mathbb{P}_Y \left((q_1^{-1}(E) \cap N) \cup (q_1^{-1}(E) \cap M)\right), \\
&= \mathbb{P}_Y \left(q_1^{-1}(E) \cap N\right) + \mathbb{P}_Y \left(q_1^{-1}(E) \cap M\right), \\
&= \mathbb{P}_Y \left(q_1^{-1}(E) \cap M\right),
\end{aligned}$$

es decir,

$$\mathbb{P}_Y \circ q_1^{-1}(E) = \mathbb{P}_Y \left(q_1^{-1}(E) \cap M\right). \tag{2.48}$$

Análogamente, se sigue que

$$\mathbb{P}_Y \circ q_2^{-1}(E) = \mathbb{P}_Y \left(q_2^{-1}(E) \cap M\right). \tag{2.49}$$

Finalmente, sea $y \in q_2^{-1}(E) \cap M$, tenemos las siguientes equivalencias:

$$\begin{aligned}
y \in q_1^{-1}(E) \cap M &\iff q_1(y) = (f(y), v_1(y)) \in E \ \wedge \ v_1(y) = v_2(y), \\
&\iff q_1(y) = (f(y), v_2(y)) \in E \ \wedge \ v_1(y) = v_2(y), \\
&\iff y \in q_2^{-1}(E) \cap M,
\end{aligned}$$

con lo cual podemos concluir que

$$q_1^{-1}(E) \cap M = q_2^{-1}(E) \cap M,$$

de donde, junto con (2.48) y (2.49), se tiene que

$$\mathbb{P}_Y \circ q_1^{-1}(E) = \mathbb{P}_Y \circ q_2^{-1}(E),$$

para cada $E \in \mathcal{F}_X \otimes \mathcal{B}(\mathbb{R})$.

Por lo tanto, hemos probado que

$$\mathbb{P}_Y \circ q_1^{-1} = \mathbb{P}_Y \circ q_2^{-1},$$

como queríamos. $\square$

Para terminar esta sección, presentamos un resultado que nos asegura la independencia de la única flecha $!_{\overline{Y}} \colon \mathbb{0} \to \overline{Y}$.

PROPOSICIÓN 2.23. *Sea $!_{\overline{Y}} \colon \mathbb{0} \to \overline{Y}$ la única flecha en **Prob** y $v \in \mathcal{L}^1(\overline{Y})$. Se tiene que v es independiente a $!_{\overline{Y}}$.*

Demostración. En primer lugar, notemos que $v^- \colon \overline{\mathbb{R}}_v \to \overline{Y}$ corresponde a la flecha en **Prob**, donde $v \colon \overline{Y} \to \overline{\mathbb{R}}$. Además, $\mathbb{0}_{!_{\overline{Y}}} = (\{*\}, \{\{*\}, \emptyset\}, \mathbb{P}_Y \circ !_Y^{-1}\})$.

Vamos a demostrar que $v \perp !_{\overline{Y}}$, para lo cual, debemos hallar $q^- \colon \mathbb{0} \otimes \overline{\mathbb{R}}_v \to \overline{Y}$, tal que el siguiente diagrama:

$$\begin{tikzcd}
\mathbb{0} \arrow[r, "!_{\overline{Y}}"] \arrow[d, "id^-_{\{*\}}"'] & \overline{Y} & \\
\mathbb{0}_{!_{\overline{Y}}} \arrow[ur, "!^{\sim}_{Y}"] \arrow[r, "p_1^-"'] & \mathbb{0}_{!_{\overline{Y}}} \otimes \overline{\mathbb{R}}_v \arrow[u, dashed, "q^-"'] & \overline{\mathbb{R}}_v \arrow[l, "p_2^-"] \arrow[ul, "v^-"']
\end{tikzcd}$$

conmuta.

Consideremos

$$\begin{aligned} q \colon Y &\longrightarrow \{*\} \times \mathbb{R} \\ y &\longmapsto (!_Y(y), v(y)) = (*, v(y)), \end{aligned}$$

con lo cual es fácil ver que el diagrama precedente conmuta. Así, basta con demostrar que q^- preserva medidas, es decir, debemos probar que

$$\mathbb{P}_Y \circ q^{-1} = \mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}.$$

Sea $E \in \{\{*\}, \emptyset\} \times \mathcal{B}(\mathbb{R})$, consideremos los siguientes casos:

- Si $E = \emptyset \times A = \emptyset$, con $A \in \mathcal{B}(\mathbb{R})$, entonces

$$
\begin{aligned}
\left(\mathbb{P}_Y \circ q^{-1}\right)(E) &= \mathbb{P}_Y(q^{-1}(E)),\\
&= \mathbb{P}_Y(q^{-1}(\emptyset)),\\
&= \mathbb{P}_Y(\emptyset),\\
&= 0.
\end{aligned}
$$

Por otro lado, tenemos que

$$
\begin{aligned}
\left(\mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}\right)(E) &= \left(\mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}\right)(\emptyset \times A),\\
&= \left(\left(\mathbb{P}_Y \circ !_Y^{-1}\right)(\emptyset)\right) \cdot \left(\left(\mathbb{P}_Y \circ v^{-1}\right)(A)\right),\\
&= \mathbb{P}_Y(!_Y^{-1}(\emptyset)) \cdot \mathbb{P}_Y(v^{-1}(A)),\\
&= \mathbb{P}_Y(\emptyset) \cdot \mathbb{P}_Y(v^{-1}(A)),\\
&= 0.
\end{aligned}
$$

Por lo tanto, tenemos que

$$
\left(\mathbb{P}_Y \circ q^{-1}\right)(E) = \mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}(E).
$$

- Ahora, supongamos que $E = \{*\} \times A = \emptyset$, con $A \in \mathcal{B}(\mathbb{R})$, tenemos que

$$
\left(\mathbb{P}_Y \circ q^{-1}\right)(E) = \mathbb{P}_Z(q^{-1}(\{*\} \times A)). \tag{2.50}
$$

Sea $y \in q^{-1}(\{*\} \times A)$, se tienen las siguientes equivalencias:

$$
\begin{aligned}
y \in q^{-1}(\{*\} \times A &\Longleftrightarrow q(y) \in \{*\} \times A,\\
&\Longleftrightarrow (!_Y(y), v(y)) \in \{*\} \times A,\\
&\Longleftrightarrow !_Y(y) \in \{*\} \wedge v(y) \in A,\\
&\Longleftrightarrow y \in !_Y^{-1}(\{*\}) \wedge y \in v^{-1}(A),\\
&\Longleftrightarrow y \in Y \wedge y \in v^{-1}(A),\\
&\Longleftrightarrow y \in v^{-1}(A),
\end{aligned}
$$

con lo cual, podemos concluir que

$$
q^{-1}(\{*\} \times A) = v^{-1}(A),
$$

de donde, junto con (2.50), obtenemos que

$$\left(\mathbb{P}_Y \circ q^{-1}\right)(E) = \mathbb{P}_Z(v^{-1}(A)).$$

Por otro lado, se tiene que

$$\begin{aligned}
\left(\mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}\right)(E) &= \left(\mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}\right)(\{*\} \times A), \\
&= \left(\left(\mathbb{P}_Y \circ !_Y^{-1}\right)(\{*\})\right) \cdot \left(\left(\mathbb{P}_Y \circ v^{-1}\right)(A)\right), \\
&= \mathbb{P}_Y(!_Y^{-1}(\{*\})) \cdot \mathbb{P}_Y(v^{-1}(A)), \\
&= \mathbb{P}_Y(Y) \cdot \mathbb{P}_Y(v^{-1}(A)), \\
&= \mathbb{P}_Y(v^{-1}(A)).
\end{aligned}$$

Así, tenemos que

$$\left(\mathbb{P}_Y \circ q^{-1}\right)(E) = \left(\mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1}\right)(E).$$

Con esto, podemos concluir que

$$\mathbb{P}_Y \circ q^{-1} = \mathbb{P}_Y \circ !_Y^{-1} \otimes \mathbb{P}_Y \circ v^{-1},$$

como queríamos. $\square$

Capítulo 3

Resultados, conclusiones y recomendaciones

3.1. Resultados

En esta sección consideraremos a $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ y $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ como espacios de probabilidad, a $f^- \colon \overline{X} \to \overline{Y}$ como una flecha en **Prob** y a $v \in \mathcal{L}^1(\overline{Y})$ como una variable aleatoria.

En primer lugar, presentamos resultados que nos garantizan que las nociones de esperanza condicional, medibilidad e independencia definidas en la categoría **Prob** corresponden a generalizaciones de sus contrapartes en teoría de probabilidades, cumpliendo de esta manera el objetivo general de este trabajo de integración curricular.

PROPOSICIÓN 3.1. *La esperanza condicional de v por f^- es una generalización de la esperanza condicional usual.*

Demostración. Sean $\mathcal{H} \subset \mathcal{F}_X$ una sub-$\sigma-$álgebra de $\mathcal{F}$ y $A \in \mathcal{H}$. Consideremos $v \in \mathcal{L}(\overline{X})$ y la flecha en **Prob** dada por $id_X^- \colon (X, \mathcal{H}, \mathbb{P}_X) \to (X, \mathcal{F}_X, \mathbb{P}_X)$, es decir, $id_X \colon (X, \mathcal{F}_X, \mathbb{P}_X) \to (X, \mathcal{H}, \mathbb{P}_X)$ es una función nula-conservadora. En efecto, es medible dado que $\mathcal{H} \subset \mathcal{F}_X$, además, $\mathbb{P}_X$ es absolutamente continua con respecto a sí misma.

Con esto, la esperanza condicional de v por id_X^-, satisface que

$$\int_A E^{id_X^-}(v)\, d\mathbb{P}_X = \int_{id^{-1}(A)} v\, d\mathbb{P}_X, \tag{3.1}$$

para cada $A \in \mathcal{H}$.

Ahora, dado que $id^{-1}(A) = A$, junto con (3.1), se tiene que

$$\int_A E^{\overline{id_X}}(v)\, d\mathbb{P}_X = \int_A v\, d\mathbb{P}_X,$$

para cada $A \in \mathcal{H}$. Esta expresión corresponde justamente a la esperanza condicional usual, es decir, $E^{\overline{id_X}}(v) = \mathbb{E}[v|\mathcal{H}]$ casi seguramente. $\square$

Lema 3.2. *El teorema 2.17 es una generalización de la siguiente propiedad que satisface la esperanza condicional usual:*

Sean $\mathcal{G} \subset \mathcal{F}_X$ una sub-$\sigma-$álgebra de $\mathcal{F}_X$, $u \in \mathcal{L}^1(\overline{X})$ y v una variable aleatoria $\mathcal{G}-$medible, se cumple que

$$\mathbb{E}[u \cdot v|\mathcal{G}] \sim_{\mathbb{P}_X} v\mathbb{E}[u|\mathcal{G}].$$

Demostración. En efecto, consideremos $\overline{id_X} \colon (X, \mathcal{G}, \mathbb{P}_X) \to (X, \mathcal{F}_X, \mathbb{P}_X)$ una flecha en **Prob**, dado que $v \sim_{\mathbb{P}_Y} v \circ id_X = v$, en particular, del teorema 2.17, tenemos que

$$E^{\overline{id_X}}(v \cdot u) \sim_{\mathbb{P}_X} vE^{\overline{id_X}}(u),$$

de donde, por la proposición 3.1, se sigue que

$$\mathbb{E}[v \cdot u|\mathcal{G}] \sim_{\mathbb{P}_X} v\mathbb{E}[u|\mathcal{G}],$$

como queríamos. $\square$

De igual forma que en la categoría **mpProb**, la siguiente proposición establece que la noción de independencia definida en la categoría **Prob** coincide exactamente con la definición de independencia clásica empleada en teoría de probabilidades.

Proposición 3.3. *Sean $\overline{X}$, $\overline{Y}$ y $\overline{Z}$ espacios de probabilidad, $f^- \colon \overline{X} \to \overline{Z}$ y $g^- \colon \overline{Y} \to \overline{Z}$ dos flechas en **Prob**. Se tiene que f^- y g^- son independientes en **Prob**, si y solo si, para cada par $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, se tiene que*

$$\mathbb{P}_Z\left(f^{-1}(A) \cap g^{-1}(B)\right) = \mathbb{P}_Z\left(f^{-1}(A)\right) \times \mathbb{P}_Z\left(g^{-1}(B)\right).$$

Demostración.

- Para la primera implicación, supongamos que las flechas f^- y g^- son independientes en **Prob**, así sabemos existe una flecha $q^- : \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \to \overline{Z}$ que preserva medidas tal que el siguiente diagrama:

$$\begin{tikzcd}
\overline{X} \arrow[r, "f^-"] \arrow[d, "id_X^-"'] & \overline{Z} & \overline{Y} \arrow[l, "g^-"'] \arrow[d, "id_Y^-"] \\
\overline{X}_f \arrow[r, "p_1^-"'] \arrow[ur, "f^\sim"] & \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \arrow[u, dashed, "q^-"'] & \overline{Y}_g \arrow[l, "p_2^-"] \arrow[ul, "g^\sim"']
\end{tikzcd}$$

conmuta, es decir,

$$q^- \circ p_1^- = f^\sim \quad \text{y} \quad q^- \circ p_2^- = g^\sim,$$

lo cual es equivalente a tener que

$$q^- \circ p_1^- = f^- \quad \text{y} \quad q^- \circ p_2^- = g^-, \tag{3.2}$$

Vamos a demostrar que

$$\mathbb{P}_Z \left(f^{-1}(A) \cap g^{-1}(B)\right) = \mathbb{P}_Z \left(f^{-1}(A)\right) \times \mathbb{P}_Z \left(g^{-1}(B)\right),$$

para cada $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$.

Sean $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, gracias a (3.2) sabemos que la flecha $q^- : \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \to \overline{Z}$ corresponde a la función q definida por

$$\begin{aligned} q : \overline{Z} &\longrightarrow \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \\ \omega &\longmapsto (f(\omega), g(\omega)). \end{aligned}$$

Con esto, tenemos que

$$\begin{aligned} \mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) &= \mathbb{P}_Z \left((q^{-1}(p_1^{-1})(A)) \cap (q^{-1}(p_2^{-1}))(B)\right), \\ &= \mathbb{P}_Z \left(q^{-1}(A \times Y) \cap q^{-1}(X \times B)\right), \\ &= \mathbb{P}_Z \left(q^{-1}\left((A \times Y) \cap (X \times B)\right)\right), \\ &= \mathbb{P}_Z \left(q^{-1}\left(A \times B\right)\right), \end{aligned}$$

$$= \left(\mathbb{P}_Z \circ q^{-1}\right)(A \times B). \tag{3.3}$$

Por otro lado, dado que $q^{-}: \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \to \overline{Z}$ preserva medidas, sabemos que

$$\mathbb{P}_Z \circ q^{-1} = \mathbb{P}_{\overline{X}_{f^-}} \otimes \mathbb{P}_{\overline{Y}_{g^-}},$$

de donde, junto con (3.3), se sigue que

$$\begin{aligned}
\mathbb{P}_Z(f^{-1}(A) \cap g^{-1}(B)) &= \mathbb{P}_{\overline{X}_{f^-}} \otimes \mathbb{P}_{\overline{Y}_{g^-}}(A \times B), \\
&= \mathbb{P}_{\overline{X}_{f^-}}(A) \cdot \mathbb{P}_{\overline{Y}_{g^-}}(B), \\
&= \left(\mathbb{P}_Z \circ f^{-1}\right)(A) \left(\mathbb{P}_Z \circ g^{-1}\right)(B), \\
&= \mathbb{P}_Z\left(f^{-1}(A)\right) \mathbb{P}_Z\left(g^{-1}(B)\right),
\end{aligned}$$

como queríamos.

- Para el recíproco supongamos que, para cada par $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, se tiene que

$$\mathbb{P}_Z\left(f^{-1}(A) \cap g^{-1}(B)\right) = \mathbb{P}_Z\left(f^{-1}(A)\right) \times \mathbb{P}_Z\left(g^{-1}(B)\right).$$

Vamos a demostrar que existe $q^{-}: \overline{X}_{f^-} \otimes \overline{Y}_{g^-} \to \overline{Z}$ que preserva medidas tal que

$$q^{-} \circ p_1^{-} = f^{\sim} \quad \text{y} \quad q^{-} \circ p_2^{-} = g^{\sim},$$

lo cual es equivalente a tener que

$$q^{-} \circ p_1^{-} = (p_1 \circ q) = f^{-} \quad \text{y} \quad q^{-} \circ p_2^{-} = (p_2 \circ q) = g^{-}.$$

Consideremos la función

$$\begin{aligned}
q \colon \overline{Z} &\longrightarrow (\overline{X}_{f^-} \otimes \overline{Y}_{g^-}) \\
\omega &\longmapsto (f(\omega), g(\omega)),
\end{aligned}$$

la cual es medible, y además se cumple que

$$p_1 \circ q = f \quad \text{y} \quad p_2 \circ q = g,$$

de donde, podemos concluir que

$$q^- \circ p_1^- = f^- \quad \text{y} \quad q^- \circ p_2^- = g^-.$$

Así, restaría con demostrar que $q\colon \overline{Z} \to \overline{X} \otimes \overline{Y}$ preserva medidas, para lo cual, debemos probar que $\mathbb{P}_Z \circ q^{-1} = \mathbb{P}_{\overline{X}_{f^-}} \otimes \mathbb{P}_{\overline{Y}_{g^-}}$. En efecto, sean $A \in \mathcal{F}_X$ y $B \in \mathcal{F}_Y$, por hipótesis, tenemos que

$$\begin{aligned}
\mathbb{P}_Z\left(q^{-1}(A \times B)\right) &= \mathbb{P}_Z\left(f^{-1}(A) \cap g^{-1}(B)\right), \\
&= \mathbb{P}_Z(f^{-1}(A))\mathbb{P}_Z(g^{-1}(B)), \\
&= \mathbb{P}_{\overline{X}_{f^-}}(A) \cdot \mathbb{P}_{\overline{Y}_{g^-}}(B), \\
&= \mathbb{P}_{\overline{X}_{f^-}} \otimes \mathbb{P}_{\overline{Y}_{g^-}}(A \times B),
\end{aligned}$$

con esto, podemos concluir que

$$\mathbb{P}_Z \circ q^{-1} = \mathbb{P}_{\overline{X}_{f^-}} \otimes \mathbb{P}_{\overline{Y}_{g^-}},$$

como queríamos. □

Ahora, presentaremos una proposición que establece que la f^-−medibilidad es una extensión de la medibilidad clásica.

Proposición 3.4. *Sean $f\colon \overline{X} \to \overline{Y}$ una flecha en **Prob** y $v \in \mathcal{L}^\infty(\overline{Y})$. Se tiene que v es f^-−medible si y solo si v es $f^{-1}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})$−medible.*

Demostración.

- Para la primera implicación, supongamos que v es f^-−medible, vamos a demostrar que v es $f^{-1}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})$−medible.

  Por hipótesis, sabemos existe $w \in \mathcal{L}^\infty(\overline{X})$ tal que $v \sim_{\mathbb{P}_Y} w \circ f$.

  Ahora, sabemos que $f^{-1}(\mathcal{F}_X) = \{\{\omega : f(w) \in B\} : B \in \mathcal{F}_X\}$ es la σ−álgebra más pequeña tal que f es medible. Luego, dado que la composición de funciones medibles es medible, obtenemos que

  $$w \circ f\colon (X, \mathcal{F}_X) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

  es $f^{-1}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})$−medible.

- Para el recíproco, supongamos que v es $f^{-1}(\mathcal{F}_X)/\mathcal{B}(\mathbb{R})$−medible, vamos a demostrar que existe $w \in \mathcal{L}^\infty(\overline{X})$ tal que $v \sim_{\mathbb{P}_Y} w \circ f$.

En efecto, basta aplicar el Lema A.3, con lo cual existe $\varphi\colon (X,\mathcal{F}_X) \to (\mathbb{R},\mathcal{B}(\mathbb{R}))$ una función $(\mathcal{F}_X/\mathcal{B}(\mathbb{R}))-$medible tal que

$$v = \varphi \circ f,$$

de donde, tomando $w = \varphi$, se sigue el resultado. □

Finalmente, presentamos resultados que se obtuvieron al estudiar aquellas categorías que tienen por objetos a los espacios de probabilidad.

PROPOSICIÓN 3.5. *La categoría **mpProb** no posee productos.*

Demostración. Consideremos la función diagonal $\pi\colon X \to X \times X$, por un lado, sean $A, B \in X$, tenemos que

$$\begin{aligned} \mathbb{P}_{X\times X}(A \times B) &= \mathbb{P}_X \otimes \mathbb{P}_X(A \times B), \\ &= \mathbb{P}_X(A)\mathbb{P}_X(B). \end{aligned} \tag{3.4}$$

Por otro lado, se tiene que

$$\begin{aligned} \mathbb{P}_X \circ \pi^{-1}(A \times B) &= \mathbb{P}_X\left(\pi^{-1}(A \times B)\right), \\ &= \mathbb{P}_X(A \cap B). \end{aligned} \tag{3.5}$$

Así, junto con (3.4) y (3.5), podemos concluir que

$$\mathbb{P}_X \circ \pi^{-1} \neq \mathbb{P}_{X\times X}.$$

Es decir, hemos probado que $\pi\colon X \to X \times X$ no es una flecha en **mpProb**. □

Consideremos la categoría **Meas**$_1$ cuyos objetos son los espacios medidos y cuyas flechas entre espacios medidos corresponden a las funciones medibles.

PROPOSICIÓN 3.6. *No es posible definir el funtor de completación en la categoría **Meas**$_1$.*

Demostración. Es suficiente probar que si $(X,\mathcal{F}_X,\mu)$, $(Y,\mathcal{F}_Y,\nu)$ son espacios medidos y $f\colon (X,\mathcal{F}_X) \to (Y,\mathcal{F}_Y)$ es $\mathcal{F}_X/\mathcal{F}_Y-$medible, entonces f no necesariamente es $\mathcal{F}_X^*/\mathcal{F}_Y^*-$medible, donde F_X^* y F_Y^* corresponden a la completación de $\mathcal{F}_X$ y $\mathcal{F}_Y$, respectivamente.

Sean $(X, \mathcal{F}_X, \mu)$ un espacio medido completo y $A \subseteq X$ tal que A no es $\mathcal{F}_X-$medible. Además, consideremos el espacio medido $(Y, \mathcal{F}_Y, \nu)$, donde $Y = \{1, 2, 3\}$, $\mathcal{F}_Y = \{\emptyset, \{1, 2, 3\}, \{1, 2\}, \{3\}\}$ y ν es la medida delta de Dirac en el punto $3 \in Y$. Definamos la función $f\colon X \to Y$ por

$$f(x) = \begin{cases} 1 & \text{si} \quad x \in A, \\ 2 & \text{si} \quad x \notin A. \end{cases}$$

Notemos que f es $\mathcal{F}_X/\mathcal{F}_Y-$medible. En efecto, sea $B \in \mathcal{F}_Y$, se tienen los siguiente casos:

1. Si $B = \emptyset$ o $B = \{3\}$, entonces $f^{-1}(B) = \emptyset$.
2. Si $B = \{1, 2, 3\}$ o $B = \{1, 2\}$, entonces $f^{-1}(B) = X$.

Así, podemos concluir que $f^{-1}(B) \in \mathcal{F}_X$.

Ahora, notemos que $\nu(\{1, 2\}) = 0$, de donde, $\{1\} \in \mathcal{F}_Y^*$ y por definición de f, se tiene que

$$f^{-1}(\{1\}) = A.$$

Por otro lado, por hipótesis tenemos $\mathcal{F}_X^* = \mathcal{F}_X$, de donde, junto con la identidad precedente, se sigue que

$$f^{-1}(\{1\}) \notin \mathcal{F}_X^*.$$

Por lo tanto, podemos concluir que f no es $\mathcal{F}_X^*/\mathcal{F}_Y^*-$medible.

$\square$

Al igual que en **CPS** podemos definir el funtor de olvido en la categoría **Prob**$^{\mathbf{op}}$ como sigue:

DEFINICIÓN 3.1 (Funtor **U**). *Sean* $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ *y* $\overline{Y} = (Y, \mathcal{F}_Y, \mathbb{P}_Y)$ *dos espacios de probabilidad. El funtor **U**:* ***Prob***$^{\mathbf{op}}$ $\to$ ***Meas*** *está definido por:*

1. *Para objetos:*

$$\begin{aligned} \boldsymbol{U}\colon \boldsymbol{Ob}(\boldsymbol{Prob}^{\mathbf{op}}) &\longrightarrow \boldsymbol{Ob}(\boldsymbol{Meas}) \\ \overline{X} &\longmapsto (X, \mathcal{F}_X), \end{aligned}$$

es decir, ***U*** *asigna un objeto* $\overline{X} = (X, \mathcal{F}_X, \mathbb{P}_X)$ *de* ***Prob*** *a* $(X, \mathcal{F}_X)$ *un objeto en* ***Meas*** *olvidando su medida de probabilidad.*

2. *Para flechas:*

$$\begin{aligned} \boldsymbol{U}\colon \boldsymbol{Prob}^{\boldsymbol{op}}(\overline{X}, \overline{Y}) &\longrightarrow \boldsymbol{Meas}(\boldsymbol{U}(\overline{X}), \boldsymbol{U}(\overline{Y})) = \boldsymbol{Meas}((X, \mathcal{F}_X), (Y, \mathcal{F}_Y)) \\ [f^-_{\text{op}}\colon \overline{X} \to \overline{Y}] &\longmapsto [f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)]. \end{aligned}$$

El funtor ***U*** *se denomina* funtor de olvido.

Observación. Notemos que el funtor **U** está bien definido en objetos y flechas. En efecto, $f^-_{\text{op}}\colon \overline{X} \to \overline{Y}$ es una flecha en **Prob**$^{\mathbf{op}}$ si $f^-\colon \overline{Y} \to \overline{X}$ es una flecha en **Prob**, de donde, por definición:

$$f\colon \overline{X} \to \overline{Y},$$

es una función $\mathcal{F}_X/\mathcal{F}_Y-$medible.

3.2. Conclusiones

- La proposición 3.1 nos permite concluir que la esperanza condicional de v por f^- es una generalización de la esperanza condicional usual.

- Gracias a la proposición 3.3 podemos asegurar que la noción de independencia definida en la categoría **Prob** es una generalización de la independencia clásica empleada en teoría de probabilidades.

- De igual forma, la proposición 3.4 nos permite inferir que la f^- medibilidad es una extensión de la medibilidad clásica.

- En la propiedad de convergencia monótona 2.14 es necesario exigir que la sucesión $(v_n)_{n\in\mathbb{N}}$ de variables aleatorias sea no negativa, dado que el teorema de convergencia monótona no se cumple para el caso de las medidas con signo.

- En [13] se demuestra el lema 2.19 aplicando el teorema 2.17, sin embargo, se puede probar este lema haciendo uso de la fórmula del cambio de variable A.4 como se realizó en este trabajo de integración curricular.

- La condición impuesta sobre las flechas de la categoría **Prob** son adecuadas, por ello, esta categoría puede ser empleada como una base para la teoría de probabilidades.
- La categoría **Prob** de espacios de probabilidad es una categoría más general que otras categorías con objetos como espacios de probabilidad.
- Las generalizaciones de la medibilidad e independencia en la Categoría **Prob** dependen de una flecha en **Prob**, mientras que sus contrapartes en teoría de probabilidades dependen de una $\sigma-$álgebra.

3.3. Recomendaciones

Más allá de uso en la topología algebraica, para lo cual fue concebida originalmente, hoy en día la teoría de categorías se emplea en varias ramas de la matemática pura como el álgebra, la lógica y la topología. En este trabajo de integración curricular describimos una aplicación de esta teoría a la teoría de probabilidades.

- Se recomienda estudiar la teoría de categorías aplicada, sus principales conceptos son la semántica funtorial y la composicionalidad. Por un lado, sean $\mathcal{C}$ y $\mathcal{D}$ categorías, la semántica funtorial se relaciona con la idea de que un functor $F\colon \mathcal{C} \to \mathcal{D}$ puede verse como una interpretación de la categoría $\mathcal{C}$ dentro de la categoría $\mathcal{D}$. Para observar esto, pensemos en la categoría $\mathcal{C}$ como una codificación de la sintaxis, mientras que la categoría $\mathcal{D}$ nos proporciona la semántica. La sintaxis se refiere a las reglas para unir cosas y la semántica explica el significado de esas cosas.

  Por su parte, la composicionalidad establece que el significado de una expresión compleja está determinado por los significados de sus partes constituyentes más pequeñas y las reglas sobre cómo combinar dichas partes.

La teoría de categorías aplicada puede explicar como el comportamiento de una red epidemiológica está modelada por un funtor, para ello, es necesario estudiar las redes de Petri y sus propiedades.

- También, se recomienda estudiar [12], en este trabajo sobre la categoría **Prob** se generalizan los conceptos filtraciones, procesos adaptados y martingalas. Una filtración generalizada se presenta como un funtor de la categoría **LCat** a la categoría **Prob**. Por su parte, los procesos adaptados se presentan como una colección de transformaciones naturales y las martingalas se presentan como un proceso adaptado haciendo uso del funtor esperanza condicional.

- Sobre la categoría **Prob** resulta interesante encontrar límites y colímites, más específicamente, determinar para qué diagramas existen tales límites, y en caso de existir, establecer cuáles de estos son preservados por el funtor esperanza condicional y el funtor de completación.

Capítulo A

Anexos

En este capítulo presentamos la demostración de ciertos teoremas y proposiciones importantes empleados en este trabajo de integración curricular, como la demostración del teorema de Radon-Nikodym.

A.1. Anexo I

Las medidas con signo permiten descomponer el espacio X en la unión de dos conjuntos medibles disjuntos, estos conjuntos se denominan conjunto positivo y conjunto negativo.

Formalizamos estas nociones con la siguiente definición:

DEFINICIÓN A.1 (Conjuntos positivos y negativos)**.** *Sean* $(X, \mathcal{F})$ *un espacio medido y* μ *una medida con signo definida sobre* $(X, \mathcal{F})$*. Un subconjunto* $F \subseteq X$*, se dice:*

- Positivo*: Si* $F \in \mathcal{F}$ *y, para todo subconjunto medible* $G \subseteq F$*, se verifica que* $\mu(G) \geq 0$*.*
- Negativo: *Si* $F \in \mathcal{F}$ *y, para todo subconjunto medible* $G \subseteq F$*, se verifica que* $\mu(G) \leq 0$*.*

Así, podemos presentar el teorema de descomposición de Hahn.

Teorema A.1 (Teorema de Descomposición de Hahn)**.** *Sean* $(X, \mathcal{F})$ *un espacio medible y* μ *una medida con signo definida sobre* $(X, \mathcal{F})$*. Existen dos subconjuntos disjuntos* P *y* N *de* X *tales que* P *es un conjunto positivo de* μ *y* N *es un conjunto negativo de* μ*, además, se satisface que*

$$X = P \cup N.$$

Se dice que el par (P, N) *es una descomposición de Hanh de* X *asociada a la medida con signo* μ*.*

Observación. Sean (P_1, N_1) y (P_2, N_2) dos descomposiciones de Hanh de X con respecto a la medida μ, consideremos el conjunto $P_1 \cap N_2$, el cual es un conjunto positivo y negativo a la vez. En efecto, notemos que

$$P_1 \cap N_2 \subset P_1 \quad P_1 \cap N_2 \subset N_2,$$

de donde, por la monotonía de μ, y la definición de conjunto positivo y conjunto negativo, respectivamente, se tiene que

$$\mu(P_1 \cap N_2) \geq 0 \quad \mu(P_1 \cap N_2) \leq 0,$$

con lo cual, podemos concluir que

$$\mu(P_1 \cap N_2) = 0.$$

De igual forma, tenemos que $P_2 \cap N_1$ es un conjunto de medida nula. Así, obtenemos que la descomposición de Hahn es esencialmente única.

Demostración del teorema de descomposición de Jordan 1.1. Por el teorema A.1, consideremos (P, N) una descomposición de Hahn de X con respecto a la medida con signo μ, es decir, P es un conjunto positivo, N es un conjunto negativo y se cumple que

$$X = P \cup N \quad \text{y} \quad P \cap N = \emptyset.$$

Con esto, sea $F \in \mathcal{F}_X$, dado que P es un conjunto positivo y $F \cap P \subset P$, tenemos que

$$V(F \cap P) \geq 0. \tag{A.1}$$

Análogamente, por definición de conjunto negativo aplicada a $F \cap N \subset N$, sabemos que

$$V(F \cap N) \leq 0. \tag{A.2}$$

Gracias a (A.1) y (A.2), podemos considerar las medidas μ^+ y μ^-, definidas por

$$\begin{aligned} \mu^+ : \mathcal{F} &\longrightarrow \overline{R}_+ \\ F &\longmapsto \mu^+(F) = \mu(F \cap P) \end{aligned} \quad \text{y} \quad \begin{aligned} \mu^- : \mathcal{F} &\longrightarrow \overline{R}_+ \\ F &\longmapsto \mu^-(F) = -\mu(F \cap N) \end{aligned}$$

Es fácil ver que μ^+ y μ^- son medidas. En efecto, para μ^+, tenemos que

- $\mu^+(\emptyset) = \mu(\emptyset \cap P) = 0$;
- Sea $\{F_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ una colección de conjuntos disjuntos dos a dos, se tiene que

$$\begin{aligned} \mu^+\left(\bigcup_{n \in \mathbb{N}} F_n\right) &= \mu\left(\bigcup_{n \in \mathbb{N}} F_n \cap P\right), \\ &= \mu\left(\bigcup_{n \in \mathbb{N}} (F_n \cap P)\right), \\ &= \sum_{n \in \mathbb{N}} \mu\left(F_n \cap P\right), \\ &= \sum_{n \in \mathbb{N}} \mu^+(F_n). \end{aligned}$$

De igual forma se prueba la $\sigma-$aditividad para μ^-.

Por otro lado, probemos que μ^+ y μ^- son mutuamente singulares, para ello, debemos probar que existen dos conjuntos F, G de X, que satisfacen la definición 1.4. Tomemos $F := P$ y $G := N$, así, tenemos que $X = F \cup G$ y $F \cap G = \emptyset$. Además, para $E \in \mathcal{F}$, se tiene que

- $\mu^*(E \cap N) = \mu((E \cap N) \cap P) = \mu(E \cap (N \cap P)) = \mu(\emptyset) = 0$;
- $\mu^*(E \cap P) = \mu((E \cap P) \cap N) = \mu(E \cap (P \cap N)) = \mu(\emptyset) = 0$,

con lo cual se sigue el resultado.

Ahora, vamos a demostrar que $\mu = \mu^+ - \mu^-$, sea $E \in \mathcal{F}_X$, dado que

$X = P \cup N$ y $P \cap N = \emptyset$, se sigue que

$$\begin{aligned}
\mu(E) &= \mu(E \cap X), \\
&= \mu(E \cap (P \cup N)), \\
&= \mu((E \cap P) \cup (E \cap N)), \\
&= \mu(E \cap P) + \mu(E \cap P), \\
&= \mu^+(E) - \mu^-(E),
\end{aligned}$$

de donde, podemos concluir que

$$\mu = \mu^+ - \mu^-,$$

como queríamos.

Finalmente, probemos que el par (μ^+, μ^-) es único en el sentido de que no depende de la descomposición de Hahn empleada. Para ello, sea (P', N') otra descomposición de Hahn de X con respecto a μ, vamos a demostrar que

$$\mu(E \cap P) = \mu(E \cap P') \quad \text{y} \quad \mu(E \cap N) = \mu(E \cap N').$$

Notemos que $E \cap (P \setminus P') = E \cap (P \cap N') \subset P$ y $E \cap (P \setminus P') \subset N'$, de donde, por definición de conjunto positivo y conjunto negativo, respectivamente, tenemos que

$$\mu(E \cap (P \setminus P')) \geq 0 \quad \text{y} \quad \mu(E \cap (P \setminus P')) \leq 0,$$

es decir,

$$\mu(E \cap (P \setminus P')) = 0. \tag{A.3}$$

Por simetría, obtenemos que

$$\mu(E \cap (P' \setminus P)) = 0. \tag{A.4}$$

Ahora, probemos que $P \cup P' = P \cup (P' \setminus P)$, en efecto, se tiene que

$$P \cup (P' \setminus P) = P \cup (P' \cap N),$$

$$\begin{aligned}
&= (P \cup P') \cap (P \cup N), \\
&= (P \cup P') \cap X, \\
&= P \cup P'.
\end{aligned}$$

Así, junto con (A.3), se sigue que

$$\begin{aligned}
\mu(E \cap (P \cup P')) &= \mu(E \cap (P \cup (P' \setminus P))), \\
&= \mu((E \cap P) \cup (E \cap (P' \setminus P))), \\
&= \mu((E \cap P)) + \mu(E \cap (P' \setminus P)), \\
&= \mu((E \cap P)).
\end{aligned}$$

Por otro lado, dado que $P' \cup P = P' \cup (P \setminus P')$ y junto con (A.4), análogamente al proceso anterior, se puede probar que

$$\mu(E \cap (P \cup P')) = \mu((E \cap P')),$$

de donde, se sigue que

$$\mu((E \cap P)) = \mu((E \cap P')),$$

para todo $E \in \mathcal{F}_X$.

De igual forma, se demuestra que

$$\mu(E \cap N) = \mu(E \cap N') = \mu(E \cap (N \cup N')). \qquad \square$$

Demostración de la Proposición 1.2. Sea $A \in \mathcal{F}$ tal que $|\nu|(A) = 0$, vamos a demostrar que $u^+(A) = 0$ y que $u^-(A) = 0$.

Usando el hecho de que $A \cap P \subset P$ y que $\mu \ll \nu$, tenemos que

$$0 \leq \mu^+(A) = \mu(A \cap P) \leq \mu(A) = |\mu|(A) = 0,$$

con lo cual, se sigue que

$$\mu^+(A) = 0.$$

Análogamente, se prueba que

$$\mu^-(A) = \mu(A \cap N) = 0.$$

Con esto, podemos concluir que

$$\mu^+ \ll \nu \quad \text{y} \quad \mu^- \ll \nu,$$

como queríamos. $\square$

A.2. Anexo II

Observación. Sea $\mathcal{C} \subseteq \mathcal{P}(X)$ un conjunto de partes de X. Si $\sigma(\mathcal{C}) = \mathcal{F}$, entonces se dice que $\mathcal{C}$ es un generador de $\mathcal{F}$.

El siguiente lema nos brinda una caracterización útil para probar la medibilidad de una función empleando generadores de $\sigma-$álgebras:

Lema A.2 (Caracterización de la medibilidad)**.** *Sean $(X, \mathcal{F}_X)$ y $(Y, \mathcal{F}_Y)$ espacios medibles. Si $\mathcal{F}_Y = \sigma(\mathcal{G})$ es generada por una colección $\mathcal{G} \subset \mathcal{P}(Y)$, entonces la función $f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ es $\mathcal{F}_X/\mathcal{F}_Y-$medible, si y solo si, para todo $G \in \mathcal{G}$, $f^{-1}(G) \in \mathcal{F}_X$.*

Demostración del Lema A.2. Consideremos la colección $\mathcal{M} := \{B \in Y : f^{-1}(B) \in \mathcal{F}_X\}$ de subconjuntos de Y, probemos que $\mathcal{M}$ es una $\sigma-$álgebra sobre $\mathcal{F}_Y$. Para lo cual, debemos probar que

1. $Y \in \mathcal{M}$.

   En efecto, dado que $f^{-1}(Y) = X \in \mathcal{M}$, se sigue que $Y \in \mathcal{M}$.

2. Si $F \in \mathcal{M}$, entonces $F^C \in \mathcal{M}$.

   Tenemos que $f^{-1}(F) \in \mathcal{F}_X$, de donde, $\left(f^{-1}(F)\right)^C = f^{-1}\left(F^C\right) \in \mathcal{F}_X$. Así, $F^C \in \mathcal{M}$.

3. Si $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$, entonces $\bigcup\limits_{n\in\mathbb{N}} F_n \in \mathcal{M}$.

   Sea $n \in \mathbb{N}$, sabemos que $F_n \in \mathcal{M}$, lo cual es equivalente a tener que $f^{-1}(F_n) \in \mathcal{F}_X$, con esto, se sigue que

$$\bigcup_{n\in\mathbb{N}} f^{-1}(F_n) = f^{-1}\left(\bigcup_{n\in\mathbb{N}} F_n\right) \in \mathcal{F}_X,$$

con lo cual podemos concluir que $\bigcup_{n\in\mathbb{N}} F_n \in \mathcal{M}$.

Por lo tanto, hemos probado que $\mathcal{M}$ es una $\sigma-$álgebra sobre $\mathcal{F}_Y$.

Ahora, por definición de $\mathcal{M}$ se tiene que $f^{-1}(G) \in \mathcal{F}_X$ para todo $G \in \mathcal{G}$, es decir, $\mathcal{G} \subset \mathcal{M}$, de donde,

$$\mathcal{F}_Y = \sigma(\mathcal{G}) \subset \mathcal{M}.$$

Así, sea $F_Y \in \mathcal{F}_Y$, tenemos que $F_Y \in \mathcal{M}$ y por definición de $\mathcal{M}$ se sigue que $f^{-1}(F_Y) \in \mathcal{F}_X$, es decir, la función $f\colon (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ es $\mathcal{F}_X/\mathcal{F}_Y-$medible. $\square$

En topología, se define la topología más débil que hace que todas las funciones en una familia dada sean continuas. En teoría de la medida, para comparar esta idea, nos referimos a la σ-álgebra generada por una colección de funciones en X, de modo que estas funciones sean medibles.

Definición A.2. *Sea I un conjunto de índices, si tenemos una colección $\{f_i : i \in I\}$ de funciones $f_i\colon X \to \mathbb{R}$, entonces*

$$\mathcal{Y} := \sigma\left(\{f_i : i \in I\}\right),$$

se define como la $\sigma-$álgebra más pequeña sobre X, tal que, para cada $i \in I$, la función f_i es $\mathcal{Y}/\mathcal{B}(\mathbb{R})-$medible. Evidentemente, tenemos que

$$\begin{aligned}\sigma\left(\{f_i : i \in C\}\right) &= \sigma\left(\{\{x \in X : f_i(X) \in B\} : \, i \in I, \; B \in \mathcal{B}(\mathbb{R})\}\right), \\ &= \sigma\left(\{f_i^{-1}(B) : \, i \in I, \; B \in \mathcal{B}(\mathbb{R})\}\right).\end{aligned}$$

Observación. Si $f\colon (X, \mathcal{F}_X) \to (R, \mathcal{B}(\mathbb{R}))$ es una variable aleatoria, entonces, claramente $\sigma(f) \subseteq \mathcal{F}_X$.

Presentamos así, el siguiente lema cuya demostración se puede encontrar en [15], este lema nos permitirá considerar una extensión de la medibilidad en la Sección 2.1.

Lema A.3. *Sean X un conjunto no vacío, $(Y, \mathcal{F}_Y)$ un espacio medible y $g\colon X \to Y, f\colon X \to \mathbb{R}$ funciones. Consideremos la $\sigma-$álgebra más pequeña, denotada por $\sigma(g)$, en el sentido de la definición A.2. Entonces f es*

$\sigma(g)/\mathcal{B}(\mathbb{R})-$*medible si y solo si existe* φ *una función* $\mathcal{F}_Y/\mathcal{B}(\mathbb{R})-$*medible tal que*

$$f = \varphi \circ g.$$

OBSERVACIÓN. Notemos que la $\sigma-$álgebra más pequeña tal que la función $g\colon X \to Y$ es $\sigma(g)/\mathcal{F}_Y$ es medible, viene dada por

$$\sigma(g) = g^{-1}(\mathcal{F}_Y).$$

En cualquier sistema matemático es de interés estudiar aquellas funciones que conservan las propiedades estructurales del sistema invariante. Es importante mencionar que no es nuestra intención estudiar en gran detalle estas funciones, simplemente, presentaremos algunos resultados fundamentales para el desarrollo de este trabajo.

Las siguientes definiciones nos permiten obtener una nueva medida a partir de una función medible.

DEFINICIÓN A.3 (Medida inducida por una función)**.** *Sean* $(X, \mathcal{F}, \mu)$ *un espacio medido y* $f\colon X \to \overline{\mathbb{R}}_+$ *una función* $\mathcal{F}-$*medible. Se define una nueva medida sobre el espacio medible* $(X, \mathcal{F})$*, dada por*

$$\nu(A) = \int_A f \, d\mu,$$

para cada $A \in \mathcal{F}_X$*.*

DEFINICIÓN A.4 (Imagen de una medida)**.** *Sean* $(X, \mathcal{F}_X, \mu)$ *un espacio medido,* $(Y, \mathcal{F}_Y)$ *un espacio medible y* $f\colon X \to Y$ *una función* $\mathcal{F}_X/\mathcal{F}_Y-$*medible. Se define la* imagen de la medida μ *por* f*, denotada por* $\mu \circ f^{-1}$*, como sigue*

$$\left(\mu \circ f^{-1}\right)(G) = \mu(f^{-1}(G)),$$

para cada $G \in \mathcal{F}_Y$*.*

OBSERVACIÓN. Notemos que $\mu \circ f^{-1}$ es una medida definida sobre $(Y, \mathcal{F}_Y)$. Además, se tiene que $(\mu \circ f^{-1})(Y) = \mu(X)$, esta propiedad es la base del teorema de cambio de variable, que se menciona a continuación, y cuya demostración se puede encontrar en [4].

TEOREMA A.4 (Fórmula de cambio de variable)**.** *Sean* $(X, \mathcal{F}_X, \mu_X)$ *un espa-*

cio medido y $(Y, \mathcal{F}_Y)$ un espacio medible. Si $T\colon (X, \mathcal{F}_X, \mu_X) \to (Y, \mathcal{F}_Y)$ es una función $\mathcal{F}_X/\mathcal{F}_Y-$ medible y $g\colon (Y, \mathcal{F}_Y) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ es una función $\mathcal{F}_Y/\mathcal{B}(\overline{\mathbb{R}})$ medible, se tiene que

$$\int_Y g\, d(u \circ T^{-1}) = \int_X g \circ T\, du.$$

Observación. La fórmula de cambio de variable A.4 nos permite pasar de una integral en $(Y, \mathcal{F}_Y)$ a una en $(X, \mathcal{F}_X)$.

Si $G \in \mathcal{F}_Y$, también se satisface que

$$\int_G g\, d(u \circ T^{-1}) = \int_{T^{-1}(G)} g \circ T\, du.$$

A.3. Anexo III

Es esta sección definimos la estructura medible del producto de dos espacios medibles, además, presentamos el teorema de Fubini.

Sean $(X, \mathcal{F}_X)$ y $(Y, \mathcal{F}_Y)$ dos espacios medibles. Consideremos el producto cartesiano $X \times Y$ y las funciones proyecciones canónicas definidas por

$$\begin{aligned} p_1 \colon X \times Y &\longrightarrow X \\ (x, y) &\longmapsto x \end{aligned} \quad \text{y} \quad \begin{aligned} p_2 \colon X \times Y &\longrightarrow Y \\ (x, y) &\longmapsto y \end{aligned}$$

Así, tenemos la siguiente definición:

Definición A.5. *La $\sigma-$álgebra* producto *de $\mathcal{F}_X$ y $\mathcal{F}_Y$, denotada por $\mathcal{F} = \mathcal{F}_X \otimes \mathcal{F}_Y$, es la $\sigma-$álgebra definida por*

$$\mathcal{F} = \sigma(p_1, p_2).$$

Es decir, la $\sigma-$álgebra $\mathcal{F}$ está generada por conjuntos de la forma

$$p_1^{-1}(B_X) = B_X \times Y \quad \textit{y} \quad p_1^{-1}(B_Y) = X \times B_Y,$$

para cada $B_X \in \mathcal{F}_X$ y cada $B_Y \in \mathcal{F}_Y$.

Ahora, dado que disponemos de una $\sigma-$álgebra definida para el producto cartesiano $S = X \times Y$, vamos a definir la medida producto sobre la

$\sigma-$álgebra $\mathcal{F} = \mathcal{F}_X \otimes \mathcal{F}_Y$.

DEFINICIÓN A.6. *Sean* $(X, \mathcal{F}_X, \mu)$ *y* $(Y, \mathcal{F}_Y, \nu)$ *dos espacios* $\sigma-$*finitos. Existe una única medida definida sobre* $\mathcal{F}_X \otimes \mathcal{F}_Y$, *denominada* medida producto *y denotada por* $\mu \otimes \nu$ *tal que*

$$\mu \otimes \nu(A \times B) = \mu(A)\mu(B),$$

para cada $(A, B) \in \mathcal{F}_X \otimes \mathcal{F}_Y$.

OBSERVACIÓN. Gracias a las definiciones A.5 y A.6 obtenemos un espacio medido sobre el producto cartesiano $X \times Y$, denotado por $(X \times Y, \mathcal{F}_X \otimes \mathcal{F}_Y, \mu \otimes \nu)$.

La hipótesis de $\sigma-$finitud en la definición A.6 nos asegura que la medida producto es única.

Para terminar esta sección, presentamos el resultado más importante, el cual nos permite bajo ciertas hipótesis intercambiar el orden en las integrales dobles.

TEOREMA A.5 (Teorema de Fubini). *Sean* $(X, \mathcal{F}_X, \mu)$ *y* $(Y, \mathcal{F}_Y, \mu)$ *dos espacios medidos* $\sigma-$*finitos y* $f \colon X \times Y \to \mathbb{R}$ *una función* $\mu \otimes \nu-$*integrable. Además, consideremos las funciones*

$$\begin{aligned} f^y \colon X &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x,y) \end{aligned} \quad \text{y} \quad \begin{aligned} f^x \colon Y &\longrightarrow \mathbb{R} \\ y &\longmapsto f(x,y). \end{aligned}$$

Se tienen los siguientes resultados

1. *La función* f^x *es* $\mu-$*integrable* $\mu-$*casi todas partes y la función* f^y *es* $\nu-$*integrable* $\nu-$*casi todas partes.*

2. *Se satisfacen las identidades*

$$\int_{X\times Y} f(x,y)\, d(\mu\otimes\nu) = \int_Y \left(\int_X f(x,y)\, d\mu \right) d\nu = \int_X \left(\int_Y f(x,y)\, d\nu \right) d\mu.$$

La demostración del teorema de Fubini se puede encontrar en [15].

A.4. Anexo IV

En esta sección presentaremos la demostración del Corolario 1.8 del teorema de Radon-Nikodym, así como la demostración de las propiedades que satisface la derivada de Radon-Nikodym de las medidas son signo descritas en la Proposición 1.9. Finalmente, presentamos el teorema de transformación, el cual emplearemos para demostrar las propiedades que satisface el funtor esperanza condicional.

Demostración del Corolario 1.8. Consideremos (ν^+, ν^-) la descomposición de Jordan de la medida con signo ν, tenemos que

$$\nu = \nu^+ - \nu^-.$$

Ahora, sabemos que ν^+ y ν^- son medidas positivas, así, por el teorema de Radon-Nikodym 1.7, sabemos existen $f_1, f_2 \in \mathcal{L}(X, \mathcal{F}, \mu)$ tales que

$$\nu^+(F) = \int_F f_1 \, d\mu \quad \text{y} \quad \nu^-(F) = \int_F f_2 \, d\mu,$$

para todo $F \in \mathcal{F}_X$.

Con esto, para $F \in \mathcal{F}_X$, tenemos que

$$\begin{aligned} \nu(F) &= \nu^+(F) - \nu^-(F) \\ &= \int_F f_1 \, d\mu - \int_F f_2 \, d\mu \\ &= \int_F (f_1 - f_2) \, d\mu, \end{aligned}$$

de donde, tomando $f = f_1 - f_2 \in \mathcal{L}(X, \mathcal{F}_X, \mu)$ se sigue el resultado. □

Demostración del Corolario 1.9.

- Por un lado, dado que ν^+ es una medida finita, por el teorema de Radon-Nikodym 1.7, sabemos existe $\frac{d\nu^+}{d\nu}$ tal que

$$\nu^+(F) = \int_F \frac{d\nu^+}{d\mu} \, d\nu, \tag{A.5}$$

  para cada $F \in \mathcal{F}$.

  Por otro lado, por el Corolario 1.8, sabemos existe $\frac{d\nu}{d\mu}$ la derivada de

Radon-Nikodym de la medida ν con respecto a la medida μ tal que

$$\nu(F) = \int_F \frac{d\nu}{d\mu}\, d\nu, \tag{A.6}$$

para cada $F \in \mathcal{F}$.

Sea $F \in \mathcal{F}$, sabemos que $\nu^+(F) = \nu(F \cap P)$, de donde, junto con (A.6), obtenemos que

$$\begin{aligned}\nu^+(F) &= \nu(F \cap P),\\ &= \int_{F\cap P} \frac{d\nu}{d\mu}\, d\mu,\end{aligned}$$

luego, dado que $\mathbb{1}_{F\cap P} = \mathbb{1}_F \cdot \mathbb{1}_P$, se tiene que

$$\begin{aligned}\nu^+(F) &= \int_X \mathbb{1}_{F\cap P} \cdot \frac{d\nu}{d\mu}\, d\mu,\\ &= \int_X \mathbb{1}_F \cdot \mathbb{1}_P \cdot \frac{d\nu}{d\mu}\, d\mu,\\ &= \int_F \mathbb{1}_P \cdot \frac{d\nu}{d\mu}\, d\mu.\end{aligned}$$

En resumen, hemos probado que

$$\nu^+(F) = \int_F \mathbb{1}_P \cdot \frac{d\nu}{d\mu}\, d\mu,$$

para todo $F \in \mathcal{F}$, con lo cual junto con (A.5), podemos concluir que

$$\frac{d\nu^+}{d\mu} = \mathbb{1}_P \cdot \frac{d\nu}{d\mu}.$$

Análogamente, dado que

$$\nu^-(F) = -\nu(F \cap N) = \int_{F\cap N} -\frac{d\nu}{d\mu}\, d\mu = \int_F -\mathbb{1}_N \cdot \frac{d\nu}{d\mu}\, d\mu,$$

se sigue que

$$\frac{d\nu^-}{d\mu} = -\mathbb{1}_N \cdot \frac{d\nu}{d\mu}.$$

- Consideremos los siguientes conjuntos:

$$P_1 = \left\{x \in X : \frac{d\nu}{d\mu} \geq 0\right\} \quad \text{y} \quad N_1 = \left\{x \in X : \frac{d\nu}{\mu} < 0\right\}.$$

Dado que $\frac{d\nu}{d\mu}$ es $\mathcal{F}/\mathcal{B}(\mathbb{R})$-medible, se tiene que P_1 y N_1 son conjuntos $\mathcal{F}-$medibles. Además, se cumple que

$$\left(\frac{d\nu}{d\mu}\right)^+ = \max\left\{\frac{d\nu}{d\mu}, 0\right\} = \mathbb{1}_{P_1}\frac{d\nu}{d\mu} \quad \text{y} \quad \left(\frac{d\nu}{d\mu}\right)^- = \max\left\{-\frac{d\nu}{d\mu}, 0\right\} = -\mathbb{1}_{N_1}\frac{d\nu}{d\mu}, \tag{A.7}$$

casi todas partes.

Vamos a demostrar que (P_1, N_1) es una descomposición de Hahn de X con respecto a la medida con signo ν. Para ello, debemos probar que

- $P_1 \cap N_1 = \emptyset$ y $P_1 \cup N_1 = X$.

  En efecto, es fácil ver que $P_1 \cap N_1 = \emptyset$. Por otro lado, para probar que $P_1 \cup N_1 = X$, basta con demostrar que $X \subseteq P_1 \cup N_1$. Por reducción al absurdo, supongamos que existe $x \in X$ tal que $x \notin P_1 \cup N_1$, de donde, se tiene que $x \in (P_1 \cup N_1)^C = P_1^C \cap N_1^C = \emptyset$, lo cual es absurdo, así, podemos colegir que

  $$X \subseteq P_1 \cup N_1,$$

  con esto, hemos probado que

  $$P_1 \cup N_1 = X.$$

- Ahora, vamos a demostrar que P_1 es un conjunto positivo y que N_1 es un conjunto negativo.

  Sea $E \in \mathcal{F}_X$ tal que $E \subset P_1$, vamos a demostrar que $\nu(E) \geq 0$. En efecto, dado que $\frac{d\nu}{d\mu}(x) \geq 0$ para $x \in P_1$, por la monotonía de la integral, tenemos que

  $$\nu(E) = \int_E \frac{d\nu}{d\mu}\, d\mu \geq 0.$$

  Así, podemos concluir que P_1 es un conjunto positivo.

  Ahora, sea $E \in \mathcal{F}_X$ tal que $E \subset N_1$, vamos a demostrar que $\nu(E) \leq 0$. Usando el hecho de que $\frac{d\nu}{d\mu}(x) \leq 0$ para $x \in N_1$, nuevamente, por la monotonía de la integral, se tiene que

$$\nu(E) = \int_E \frac{d\nu}{d\mu}\, d\mu \leq 0.$$

Por lo tanto, hemos probado que (P_1, N_1) es una descomposición de Hahn de X con respecto a la medida con signo ν.

Así, junto con (A.7) hemos probado que

$$\left(\frac{d\nu}{d\mu}\right)^+ = \mathbb{1}_{P_1}\frac{d\nu}{d\mu} = \frac{d\nu^+}{d\mu} \quad \text{y} \quad \left(\frac{d\nu}{d\mu}\right)^- = -\mathbb{1}_{N_1}\frac{d\nu}{d\mu} = \frac{d\nu^-}{d\mu},$$

casi todas partes. □

Ahora, consideremos el siguiente resultado que nos será de utilidad en el Sección 2.1:

TEOREMA A.6 (Teorema de transformación)**.** *Sea $g\colon (Y, \mathcal{F}_Y) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ una función $\mathcal{F}_Y/\mathcal{B}(\mathbb{R})-$medible. Si T es una transformación medible de un espacio medible $(X, \mathcal{F}_X, \mu_X)$ en un espacio medible totalmente $\sigma-$finito $(Y, \mathcal{F}_Y, \mu_Y)$ tal que $u \circ T^{-1}$ es absolutamente continua con respecto μ_Y, entonces existe una función medible no negativa $\phi = \frac{d(u_X \circ T^{-1})}{du_Y}$ tal que*

$$\int g(T(x))\, d\mu_X(x) = \int g(y)\phi(y)\, d\mu_Y(y),$$

en el sentido de que si una integral existe, entonces también existe la otra y las dos son iguales.

OBSERVACIÓN. La función ϕ desempeña el papel del Jacobiano (o, más bien, el valor absoluto del Jacobiano) en la teoría de las transformaciones de múltiples integrales.

En este caso, la función ϕ corresponde a la derivada de Radon-Nikodym de la medida $\mu_x \circ T^{-1}$.

Referencias bibliográficas

[1] Takanori Adachi. Toward Categorical Risk Measure Theory. *Theory and Applications of Categories*, 29:389–405, 2014.

[2] Uwe Franz. What is Stochastic Independence? *Quantum Probability and White Noise Analysis, Non-Commutativity, Infinite-Dimensionality and Probability at the Crossroads*, pages 254–274, 2003.

[3] Michèle Giry. A categorical approach to probability theory. *Lecture Notes in Mathematics*, 915:84–102, 2016.

[4] Paul Richard Halmos. *Measure Theory*. Springer-Verlag, Estados Unidos, 1974.

[5] Bart Jacobs. From Probability Monads to Commutative Effectuses. *Logical and Algebraic Methods in Programming*, 2017.

[6] Franz Lawvere. The category of probabilistic mappings. 1962.

[7] Dan Marsden. Category Theory Using String Diagrams. *Logic in Computer Science*, 2014.

[8] Hitoshi Motoyama. Classical and quantum conditional measures from a categorical viewpoint. *Notes on the Research Institute of Mathematical Analysis*, 2059:84–102, 2017.

[9] Paolo Perrone. *Categorical Probability and Stochastic Dominance in Metric Spaces*. Max Planck Institute for Mathematics in the Sciences Leipzig, Alemania, 2018.

[10] Paolo Perrone. *Notes on Category Theory with examples from basic mathematics*. arXiv, Estados Unidos, 2019.

[11] Albert Shiryaev. *Probability-1*. Springer, Estados Unidos, 2016.

[12] Katsushi Nakajima Takanori Adachi and Yoshihiro Ryu. A binomial asset pricing model in a categorical setting. 2019.

[13] Yoshihiro Ryu Takanori Adachi. A Category of Probability Spaces. 2018.

[14] Cédric Villani. Optimal transport: old and new. *Grundlehren der mathematischen Wissenschaften. Springer*, 338, 2009.

[15] David Williams. *Probability with Martingales*. Cambridge University Press, Estados Unidos, 1991.