Probabilidad y estadística para la Inteligencia Artificial: Variables Aleatorias

Camilo Enrique Argoty Pulido

Especialización en Inteligencia Artificial

2 de septiembre de 2025

Variables aleatorias

Variables aleatorias

Una variable aleatoria es una función (medible) que permite traducir los resultados de un experimento en un espacio de probabilidad, en un experimento equivalente pero en un conjunto numérico, como los números naturales ($\mathbb{N}$) o los números reales ($\mathbb{R}$).

Variables aleatorias

Una variable aleatoria es una función (medible) que permite traducir los resultados de un experimento en un espacio de probabilidad, en un experimento equivalente pero en un conjunto numérico, como los números naturales ($\mathbb{N}$) o los números reales ($\mathbb{R}$).

Si la función toma valores en un conjunto discreto, se dice que la variable aleatoria es *discreta*:

$$X : \Omega \to \mathbb{N}$$

Variables aleatorias

Una variable aleatoria es una función (medible) que permite traducir los resultados de un experimento en un espacio de probabilidad, en un experimento equivalente pero en un conjunto numérico, como los números naturales ($\mathbb{N}$) o los números reales ($\mathbb{R}$).

Si la función toma valores en un conjunto discreto, se dice que la variable aleatoria es *discreta*:

$$X : \Omega \to \mathbb{N}$$

Si la variable toma valores en $\mathbb{R}$, se denomina *continua*:

$$X : \Omega \to \mathbb{R}$$

Variables aleatorias

Una variable aleatoria es una función (medible) que permite traducir los resultados de un experimento en un espacio de probabilidad, en un experimento equivalente pero en un conjunto numérico, como los números naturales ($\mathbb{N}$) o los números reales ($\mathbb{R}$).

Si la función toma valores en un conjunto discreto, se dice que la variable aleatoria es *discreta*:

$$X : \Omega \to \mathbb{N}$$

Si la variable toma valores en $\mathbb{R}$, se denomina *continua*:

$$X : \Omega \to \mathbb{R}$$

El beneficio de esta aproximación es que con variables aleatorias, las probabilidades se calculan por medio de sumas e integrales (En el fondo las sumas son también integrales)

Función de distribución

Dada una v.a. (variable aleatoria) X la **función de distribución** asociada a X es la función que a cada valor $x \in \mathbb{R}$ asigna la probabilidad de que X sea menor que x:

Función de distribución

Dada una v.a. (variable aleatoria) X la **función de distribución** asociada a X es la función que a cada valor $x \in \mathbb{R}$ asigna la probabilidad de que X sea menor que x:

$$F_X(x) = P(X \leq x)$$

Función de densidad de probabilidad

Dada una v.a. X, la **función de densidad** es aquella función $f_X(x)$ que permite calcular la probabilidad que X tome valores en un cierto conjunto S como la integral de f_X en dicho conjunto.

Función de densidad de probabilidad

Dada una v.a. X, la **función de densidad** es aquella función $f_X(x)$ que permite calcular la probabilidad que X tome valores en un cierto conjunto S como la integral de f_X en dicho conjunto. En otras palabras, si X es continua:

Función de densidad de probabilidad

Dada una v.a. X, la **función de densidad** es aquella función $f_X(x)$ que permite calcular la probabilidad que X tome valores en un cierto conjunto S como la integral de f_X en dicho conjunto. En otras palabras, si X es continua:

$$P(X \in S) = \int_S f_X(x)dx$$

Función de densidad de probabilidad

Dada una v.a. X, la **función de densidad** es aquella función $f_X(x)$ que permite calcular la probabilidad que X tome valores en un cierto conjunto S como la integral de f_X en dicho conjunto. En otras palabras, si X es continua:

$$P(X \in S) = \int_S f_X(x)dx$$

Si X es discreta,

Función de densidad de probabilidad

Dada una v.a. X, la **función de densidad** es aquella función $f_X(x)$ que permite calcular la probabilidad que X tome valores en un cierto conjunto S como la integral de f_X en dicho conjunto. En otras palabras, si X es continua:

$$P(X \in S) = \int_S f_X(x)dx$$

Si X es discreta,

$$P(X \in S) = \sum_{x_i \in S} p_i$$

Errores conceptuales comunes, alrededor de las funciones de distribución y densidad

Suele haber confusiones comunes en los estudiantes sobre las distribuciones y densidades, entre otras:

Errores conceptuales comunes, alrededor de las funciones de distribución y densidad

Suele haber confusiones comunes en los estudiantes sobre las distribuciones y densidades, entre otras:

- Creer que las funciones de distribución y densidad, son lo mismo.

Errores conceptuales comunes, alrededor de las funciones de distribución y densidad

Suele haber confusiones comunes en los estudiantes sobre las distribuciones y densidades, entre otras:

- Creer que las funciones de distribución y densidad, son lo mismo.
- Creer que la función de densidad es lo mismo que una probabilidad. Esto sólo es cierto para v.a. discretas.

Errores conceptuales comunes, alrededor de las funciones de distribución y densidad

Suele haber confusiones comunes en los estudiantes sobre las distribuciones y densidades, entre otras:

- Creer que las funciones de distribución y densidad, son lo mismo.
- Creer que la función de densidad es lo mismo que una probabilidad. Esto sólo es cierto para v.a. discretas.
- Por lo anterior, creer que los valores de una función de densidad son menores que 1. Las que son menores que 1 son las funciones de distribución.

Errores conceptuales comunes, alrededor de las funciones de distribución y densidad

Suele haber confusiones comunes en los estudiantes sobre las distribuciones y densidades, entre otras:

- Creer que las funciones de distribución y densidad, son lo mismo.
- Creer que la función de densidad es lo mismo que una probabilidad. Esto sólo es cierto para v.a. discretas.
- Por lo anterior, creer que los valores de una función de densidad son menores que 1. Las que son menores que 1 son las funciones de distribución.

Punto clave a recordar:

Una función de densidad es una función que nos ayuda a calcular probabilidades como una integral.

Variables aleatorias discretas comunes

Nombre	Parámetros	$p(x)$	Uso
Uniforme	$N > 0$, entero	$\frac{1}{N}$, $x = 1, \ldots, N$	Fenómenos acotados cuyo comportamiento no se entiende.
Binomial	$p \in (0, 1)$, $n \in \mathbb{N}_{\geq 0}$	$\binom{n}{x} p^x (1 - p)^{n-x}$ $x \in \{0, 1 \ldots, n\}$	Probabilidad de x éxitos y $n - x$ fracasos.
Poisson	$\lambda > 0$	$\frac{e^{-\lambda}\lambda^x}{x!}$, $x \in \mathbb{N}_{\geq 0}$	Conteo de ocurrencias de un evento durante un periodo de tiempo.
Geométrica	$p \in (0, 1)$	$(1 - p)^x p$, $x \in \mathbb{N}_{\geq 0}$	Número de fracasos hasta el primer éxito.
Binomial negativa	$r > 0$, $p \in (0, 1)$	$\binom{r+x-1}{x} p^r (1 - p)^x$, $x \in \mathbb{N}_{\geq 0}$	Número de fracasos hasta el r-ésimo éxito.

Variables aleatorias continuas comunes

Nombre	Parámetros	$f_X(x)$	Uso
Uniforme	$a < b \in \mathbb{R}$	$\frac{1}{b-a}\mathbb{I}_{(a,b)}$	Fenómenos acotados cuyo comportamiento no se entiende.
Normal	$\mu, \sigma^2 \geq 0$	$\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$	Promedios y errores. Hipótesis en muchas técnicas estadísticas y para aproximar otras distribuciones.
Exponencial	$\frac{1}{\lambda} = \theta > 0$	$\lambda e^{-\lambda x}$, $x > 0$	Tiempo entre eventos. Relacionada con la distribución Poisson.
Ji cuadrado	$k > 0$, grados de libertad	$\frac{(\frac{1}{2})^{\frac{k}{2}}}{\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$ $x > 0$	Distribución de la varianza de una normal. Suma de cuadrados de normales.
Gamma	$\alpha > 0$, $\beta > 0$	$\frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$, $x > 0$	Generaliza otras distribuciones (exponencial y Ji-cuadrado).

Funciones de distribución y de densidad empíricas en Python

Ver Notebook de Jupyter

Vectores aleatorios

Si se tiene una función (medible) de un espacio de probabilidad (Ω, p) en un espacio vectorial $\mathbb{R}^n$, se dice que se tiene un **vector aleatorio**.

Vectores aleatorios

Si se tiene una función (medible) de un espacio de probabilidad (Ω, p) en un espacio vectorial $\mathbb{R}^n$, se dice que se tiene un **vector aleatorio**.

¿Cómo se reinterpreta el concepto de densidad de probabilidad para el caso de un vector aleatorio?

Vectores aleatorios

Si se tiene una función (medible) de un espacio de probabilidad (Ω, p) en un espacio vectorial $\mathbb{R}^n$, se dice que se tiene un **vector aleatorio**.

¿Cómo se reinterpreta el concepto de densidad de probabilidad para el caso de un vector aleatorio?

Se debe recordar que la clave de una función de densidad de probabilidad es que es una función que permite calcular probabilidades como integrales.

Vectores aleatorios

Si se tiene una función (medible) de un espacio de probabilidad (Ω, p) en un espacio vectorial $\mathbb{R}^n$, se dice que se tiene un **vector aleatorio**.

¿Cómo se reinterpreta el concepto de densidad de probabilidad para el caso de un vector aleatorio?

Se debe recordar que la clave de una función de densidad de probabilidad es que es una función que permite calcular probabilidades como integrales.

En el caso de un vector aleatorio, la integral que se debe usar debe ser una integral múltiple.

Vectores aleatorios

Si se tiene una función (medible) de un espacio de probabilidad (Ω, p) en un espacio vectorial $\mathbb{R}^n$, se dice que se tiene un **vector aleatorio**.

¿Cómo se reinterpreta el concepto de densidad de probabilidad para el caso de un vector aleatorio?

Se debe recordar que la clave de una función de densidad de probabilidad es que es una función que permite calcular probabilidades como integrales.

En el caso de un vector aleatorio, la integral que se debe usar debe ser una integral múltiple.

Por lo tanto, una función de densidad para un vector aleatorio $\mathbb{X}$ con valores en $\mathbb{R}^n$, debe ser un campo escalar $f_{\mathbb{X}} : \mathbb{R}^n \to \mathbb{R}$

La densidad conjunta de varias variables aleatorias

Si se tienen n v.a. X_1, X_2,..., X_n, se puede definir un vector aleatorio $\mathbb{X} = (X_1, X_2, \ldots, X_n)$

La densidad conjunta de varias variables aleatorias

Si se tienen n v.a. X_1, X_2,..., X_n, se puede definir un vector aleatorio $\mathbb{X} = (X_1, X_2, \ldots, X_n)$

La función de densidad de este vector aleatorio, se denomina la **función de densidad conjunta** de X_1, X_2,..., X_n y se denota por $f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)$.

Ejemplos

Example

Suponga que X e Y son v.a. discretas cuya distribución conjunta es:

Ejemplos

Example

Suponga que X e Y son v.a. discretas cuya distribución conjunta es:

X

		-1	0	1
Y	1	$1/9$	$1/9$	$1/9$
	0	$1/9$	0	$1/6$
	-1	$1/6$	$1/9$	$1/9$

Ejemplos

Example

Suponga que X e Y son v.a. discretas cuya distribución conjunta es:

$$
\begin{array}{cc|ccc}
 & & & X & \\
 & & -1 & 0 & 1 \\
\hline
 & 1 & 1/9 & 1/9 & 1/9 \\
Y & 0 & 1/9 & 0 & 1/6 \\
 & -1 & 1/6 & 1/9 & 1/9
\end{array}
$$

Encuentre la probabilidad de que $Y > X$

Ejemplos

Example

Suponga que X e Y son v.a. discretas cuya distribución conjunta es:

X

$$\begin{array}{cccccc} & & -1 & 0 & 1 \\ \multirow{4}{*}{Y} & 1 & 1/9 & 1/9 & 1/9 \\ & 0 & 1/9 & 0 & 1/6 \\ & -1 & 1/6 & 1/9 & 1/9 \end{array}$$

Encuentre la probabilidad de que $Y > X$

Example

Suponga que $f_{X,Y}(x,y) = \frac{3}{2}(x^2 + y^2)$ es la función de densidan conjunta de dos variables X e Y, definida sobre el cuadrado unidad $0 \leq x, y \leq 1$.

Ejemplos

Example

Suponga que X e Y son v.a. discretas cuya distribución conjunta es:

X

		-1	0	1
Y	1	$1/9$	$1/9$	$1/9$
	0	$1/9$	0	$1/6$
	-1	$1/6$	$1/9$	$1/9$

Encuentre la probabilidad de que $Y > X$

Example

Suponga que $f_{X,Y}(x,y) = \frac{3}{2}(x^2 + y^2)$ es la función de densidan conjunta de dos variables X e Y, definida sobre el cuadrado unidad $0 \leq x, y \leq 1$.

Calcule la probabilidad de que $Y > X$

Densidades marginales

Densidades marginales y condicionales

Cuando se conoce la densidad conjunta de una tupla de v.a., las densidades de cada variable por separado se conocen como **densidades marginales**.

Densidades marginales y condicionales

Cuando se conoce la densidad conjunta de una tupla de v.a., las densidades de cada variable por separado se conocen como **densidades marginales**.

A partir de una densidad conjunta, es posible calcular las densidades marginales de las componentes del vector aleatorio.

Densidades marginales y condicionales

Cuando se conoce la densidad conjunta de una tupla de v.a., las densidades de cada variable por separado se conocen como **densidades marginales**.

A partir de una densidad conjunta, es posible calcular las densidades marginales de las componentes del vector aleatorio.

El cálculo se deriva del proceso de integración de una integral múltiple en una iteración de integrales. Así, en el caso de dos v.a. X e Y:

Densidades marginales y condicionales

Cuando se conoce la densidad conjunta de una tupla de v.a., las densidades de cada variable por separado se conocen como **densidades marginales**.

A partir de una densidad conjunta, es posible calcular las densidades marginales de las componentes del vector aleatorio.

El cálculo se deriva del proceso de integración de una integral múltiple en una iteración de integrales. Así, en el caso de dos v.a. X e Y:

$$1 = \iint_S f_{X,Y}(x,y)dA = \int_a^b f_X(x)dx$$

Densidades marginales y condicionales

Cuando se conoce la densidad conjunta de una tupla de v.a., las densidades de cada variable por separado se conocen como **densidades marginales**.

A partir de una densidad conjunta, es posible calcular las densidades marginales de las componentes del vector aleatorio.

El cálculo se deriva del proceso de integración de una integral múltiple en una iteración de integrales. Así, en el caso de dos v.a. X e Y:

$$1 = \iint_S f_{X,Y}(x,y)dA = \int_a^b f_X(x)dx$$

donde

$$f_X(x) = \int_{c(x)}^{d(x)} f_{X,Y}(x,y)dy$$

Densidades condicionales

De lo anterior,

Densidades condicionales

De lo anterior,

$$\int_{c(x)}^{d(x)} \frac{f_{X,Y}(x,y)}{f_X(x)} dy = 1$$

Densidades condicionales

De lo anterior,

$$\int_{c(x)}^{d(x)} \frac{f_{X,Y}(x,y)}{f_X(x)} dy = 1$$

Esto quiere decir que la función

Densidades condicionales

De lo anterior,

$$\int_{c(x)}^{d(x)} \frac{f_{X,Y}(x,y)}{f_X(x)} dy = 1$$

Esto quiere decir que la función

$$\frac{f_{X,Y}(x,y)}{f_X(x)}$$

Densidades condicionales

De lo anterior,

$$\int_{c(x)}^{d(x)} \frac{f_{X,Y}(x,y)}{f_X(x)} dy = 1$$

Esto quiere decir que la función

$$\frac{f_{X,Y}(x,y)}{f_X(x)}$$

Es una densidad de probabilidad en el intervalo $(c(x), d(x))$.

Densidades condicionales

De lo anterior,

$$\int_{c(x)}^{d(x)} \frac{f_{X,Y}(x,y)}{f_X(x)} dy = 1$$

Esto quiere decir que la función

$$\frac{f_{X,Y}(x,y)}{f_X(x)}$$

Es una densidad de probabilidad en el intervalo $(c(x), d(x))$.

Esta densidad se conoce como la **densidad condicional** de Y dado X, y se denota como $f_{Y|X=x}(y|x)$.

Densidades marginales: Ejemplo continuo

Sean X e Y dos v.a. El vector aleatorio (X, Y) tiene función de densidad conjunta:

Densidades marginales: Ejemplo continuo

Sean X e Y dos v.a. El vector aleatorio (X, Y) tiene función de densidad conjunta:

$$f_{(X,Y)}(x, y) = k(x^2 + y^2)$$

Densidades marginales: Ejemplo continuo

Sean X e Y dos v.a. El vector aleatorio (X, Y) tiene función de densidad conjunta:

$$f_{(X,Y)}(x, y) = k(x^2 + y^2)$$

Definida para $0 \leq x \leq 1$ y $0 \leq y \leq x$.

Densidades marginales: Ejemplo continuo

Sean X e Y dos v.a. El vector aleatorio (X, Y) tiene función de densidad conjunta:

$$f_{(X,Y)}(x, y) = k(x^2 + y^2)$$

Definida para $0 \leq x \leq 1$ y $0 \leq y \leq x$.

- ¿Cuánto vale k?

Densidades marginales: Ejemplo continuo

Sean X e Y dos v.a. El vector aleatorio (X, Y) tiene función de densidad conjunta:

$$f_{(X,Y)}(x, y) = k(x^2 + y^2)$$

Definida para $0 \leq x \leq 1$ y $0 \leq y \leq x$.

- ¿Cuánto vale k?
- Hallar la densidad marginal de X y la densidad condicional de Y dado X.