II.3 Variable Aleatoria Continua

March 31, 2019

Variable Aleatoria Continua

Recordemos que, dado el **espacio de probabilidad** $\mathscr{E}(\Omega, \mathscr{F}, P)$, definimos la **variable aleatoria** X

$$\begin{aligned} X : & \longmapsto \mathbb{R} \\ \alpha & \longmapsto X(\alpha) \end{aligned}$$

Consideremos ahora (a diferencia del caso discreto) que la variable aleatoria X toma valores sobre el conjunto de los números reales $\mathbb{R}$, *variando de forma continua la asignación de probabilidades en intervalos de la recta real*. Esto hace que, en general, las probabilidades asignadas a cada valor (pero no a cada intervalo) de la variable aleatoria sean nulas:

$$P(X = x) = 0 \quad \forall x \in \mathbb{R}$$

Es posible también que nos encontremos ante una **variable aleatoria mixta**, llamada así por compartir características de las discretas y de las continuas, en cuyo caso hay algunos $x_i \in \mathbb{R}$ con masas de probabilidad asignadas y $\sum_i P(x_i) < 1$ y el resto de la probabilidad se distribuye de forma continua en intervalos.

Función de distribución de probabilidad (FDP)

Puede demostrarse que cualquier suceso de $\mathscr{F}$ puede obtenerse mediante operaciones conjuntistas realizadas con sucesos definidos sobre semi-intervalos de la forma $\{X \leq x\}$, donde $x \in \mathbb{R}$ es un número real arbitrario. Si se tienen las probabilidades de tales sucesos se conoce totalmente el espacio de probabilidad de la variable aleatoria, lo que justifica la definición de **función de distribución de probabilidad (FDP)** como:

$$F_X(x) = P(\{X \leq x\}) \quad -\infty \leq x \leq \infty$$

El **cuantil** x_q corresponde al valor que hace que el q% de la distribución esté por debajo. Por tanto, $q = F_X(x_q) \implies x_q = F_X^{-1}(q)$ y la *función inversa* F_X^{-1} se llama **función cuantil**. Por ejemplo,

- El primer cuartil es $x_{0.25} = F_X^{-1}(0.25)$
- La mediana o segundo cuartil es $x_{0.5} = F_X^{-1}(0.5)$
- El tercer cuartil es $x_{0.75} = F_X^{-1}(0.75)$
- El octavo decil es $x_{0.8} = F_X^{-1}(0.8)$
- El décimoquinto percentil es $x_{0.15} = F_X^{-1}(0.15)$

Propiedades de la función de distribución de probabilidad

Sean $F_X(x^-) = \lim_{\epsilon \to 0} F_X(x - \epsilon)$ y $F_X(x^+) = \lim_{\epsilon \to 0} F_X(x + \epsilon)$ los valores de la función de distribución en x al acercarse respectivamente por la izquierda y por la derecha. Se cumplen las siguientes propiedades:

1. $F_X(-\infty) = 0 \quad F_X(+\infty) = 1$
2. $P(\{X > x\}) = 1 - F_X(x)$
3. F_X es continua por la derecha: $F_X(x^+) = F_X(x)$
4. Si la variable aleatoria es continua, F_X también lo es y $F_X(x^-) = F_X(x)$
5. F_X es no decreciente: $x_1 < x_2 \implies F_X(x_1) \leq F_X(x_2)$
6. $F_X(x_0) = 0 \implies F_X(x) = 0 \quad \forall x \leq x_0$
7. En los puntos de discontinuidad, $P(\{X = x\}) = F_X(x) - F_X(x^-)$
8. $P(\{x_1 < X \leq x_2\}) = F_X(x_2) - F_X(x_1)$
9. $P(\{x_1 \leq X \leq x_2\}) = F_X(x_2) - F_X(x_1^-)$
10. Si la **variable aleatoria es discreta**, la función de distribución es constante salvo en los puntos de discontinuidad en los que se sitúan los sucesos muestrales elementales.
11. Si la **variable aleatoria es mixta**, la función de distribución es continua salvo en los puntos en los que están localizadas las masas de probabilidad discreta, en los que la función presenta discontinuidades de salto igual a tales probabilidades.

Función de densidad de probabilidad (fdp)

Suele ser conveniente hacer un símil entre probabilidad y masa, de modo que nuestra distribución de probabilidad podemos interpretarla como una masa igualmente distribuida a lo largo del eje real.

Si la **variable es discreta**, tendremos **masas puntuales** ideales en puntos del eje real, mientras que si la **variable es continua**, la masa se distribuirá concentrándose más o menos a lo largo del eje, pero sin llegar a formar masas puntuales. La **variable mixta** combina ambos tipos de masa a lo largo del eje.

Como ya se ha visto, **la función de distribución describe la masa acumulada hasta una posición** y, por diferencias, puede calcularse la masa de cualquier intervalo.

La **función de densidad**, como su nombre indica, describe la densidad de masa (masa por unidad de longitud) a lo largo del eje.

La **función de densidad de probabilidad** (fdp) se define:

$$f_X(x) = \frac{dF_X(x)}{dx} = \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x/2) - F_X(x - \Delta x/2)}{\Delta x} = \lim_{\Delta x \to 0} \frac{P\left(\{(x - \Delta x/2) < x \leq (x + \Delta x/2)\}\right)}{\Delta x}$$

Si **la variable aleatoria es continua**,

$$P\left(\{(x - \Delta x/2) < x \leq (x + \Delta x/2)\}\right) \approx f_X(x)\Delta x$$

Si **la variable aleatoria es discreta**, $F_X(x)$ resulta discontinua por la izquierda en los puntos donde se localizan masas de probabilidad y no es derivable en los mismos. En tales puntos x_i puede definirse la **función de masa de probabilidad** $p_X(x_i) = P(\{X = x_i\})$, ya vista anteriormente.

Matemáticamente, puede usuarse la función $\delta(x)$, que tiene **área uno** localizada en $x = 0$, anchura nula y altura infinita para relacionar las funciones de masa y de densidad de probabilidad. Si la variable aleatoria es discreta:

$$f_X(x) = \sum_i p_X(x_i)\delta(x - x_i)$$

Propiedades de la función de densidad de probabilidad (fdp)

La función de distribución de probabilidad (FDP) se obtiene integrando:

$$f_X(x) = \frac{dF_X(x)}{dx} \iff F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

Dado que $F_X(\infty) = 1$ se concluye que $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Por tanto, el área debajo de la función de densidad de probabilidad **siempre** es uno. Esto es lógico, pues el eje real corresponde al espacio muestral o suceso seguro.

La probabilidad de un intervalo es

$$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)dx$$

y corresponde al área debajo de la curva de la fdp en dicho intervalo. Adviértase que $\int_{0^-}^{0^+} \delta(x)dx = 1$, lo que nos permite extender el formalismo a variables discretas y mixtas.

Propiedades de la función de masa de probabilidad (fmp)

Como ya se ha visto, cuando la variable aleatoria es discreta, se define la **función de masa de probabilidad**, que se relaciona como sigue con la fdp:

$$p_X(x_i) = P(\{X = x_i\}) \iff f_X(x) = \sum_i p_X(x_i)\delta(x - x_i)$$

Podemos relacionar la función de masa de probabilidad (fmp) con la función de distribución de probabilidad (FDP), que corresponde a la función acumulada de probabilidad vissta anteriormente:

$$p_X(x_n) = F_X(x_n) - F_X(x_{n-1}) \iff F_x(x_n) = \sum_{i=-\infty}^{n} p_X(x_i)$$

Siguen fácilmente las siguientes propiedades:

$$F_x(\infty) = 1 \implies \sum_{i=-\infty}^{+\infty} p_X(x_i) = 1$$

$$P(x_{n_1} \leq X \leq x_{n_2}) = \sum_{i=n_1}^{n_2} p_X(x_i)$$

Relación entre el histograma y la fdp Supongamos que tenemos una muestra estadística con N observaciones de una variable estadística. Podemos considerar que tales observaciones corresponden a realizaciones de un experimento modelado por un espacio de probabilidad sobre el que hemos definido una variable aleatoria correspondiente con la variable estadística.

El espacio de probabilidad propociona un modelo teórico de los datos, mientras que las muestra estadísitca resulta de la experiencia empírica. £Podemos caracterizar probabilísticamente la población a partir de la muetsra estadística?

Como sabemos podemos construir un histograma de frecuencias relativas, sin más que agrupar los valores de la variable estadística por intervalos. Considerando que todos los intervalos tienen igual anchura Δx, podemos aproximar la fdp en cada x_i (puntos medios del intervalo, como sigue, siendo F_i la frecuencia absoluta de observaciones que caen dentro del intervalo:

$$f_X(x_i) \approx \frac{F_i}{N\Delta x}$$

Esperanza matemática y momentos

La esperanza matemática o media de una variable aleatoria X es:

$$\eta_X = E\{X\} = \int_{-\infty}^{\infty} x f_X(x) dx$$

Si consideramos la aproximación anterior del histograma a la fdp, podemos ver que la media muestral aproxima la esperanza matemática: $\hat{x} = \sum_{i=1}^{N} x_i \frac{F_i}{\Delta x} \approx \eta_X$.

La esperanza matemática puede definirse sobre cualquier función $g(X)$ de la variable aleatoria:

$$\eta_{g(X)} = E\{g(X)\} = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

La **esperanza matemática es un operador lineal**, esto es, sean $g_1(X)$ y $g_2(X)$:

$$E\{a_1 g_1(X) + a_2 g_2(X)\} = a_1 E\{g_1(X)\} + a_2 E\{g_2(X)\}$$

Podemos así definir los **momentos**:

$$m_n = E\{X^n\} = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

En particular, la **media** es $m_1 \equiv \eta_X = E\{X^1\}$ y el **valor cuadrático medio** es $m_2 = E\{X^2\}$

Y los **momentos centrales**:

$$\mu_n = E\{(X - \eta_X)^n\} = \int_{-\infty}^{\infty} (x - \eta_X)^n f_X(x) dx$$

En particular, la **varianza** es $m_2 \equiv \sigma_X^2 = E\{(X - \eta_X)^2\}$. Aplicando la linealidad:

$$\sigma_X^2 = E\{(X - \eta_X)^2\} = E\{X^2\} - \eta_X^2$$

Esperanza y momentos en el caso discreto

Como ya sabemos, si la **variable aleatoria es discreta**, se utilizan las funciones de masa de probabilidad en vez de las fdp, y sumas en vez de integrales, repitiéndose aquí las esperanzas y momentos para el caso discreto vistos anteriormente:

$$\eta_{g(X)} = E\{g(X)\} = \sum_{i=-\infty}^{\infty} g(x_i) p_X(x_i)$$

$$m_n = E\{X^n\} = \sum_{i=-\infty}^{\infty} x^n f_X(x) dx$$

$$\mu_n = E\{(X - \eta_X)^n\} = \sum_{i=-\infty}^{\infty} (x - \eta_X)^n f_X(x) dx$$

Todos estos momentos tienen sus equivalentes muestrales, ya vistos en la parte de introducción a la estadística.

Por otra parte, es muy habitual **normalizar** una variable aleatoria X, construyendo otra $Y = \frac{X-\eta_X}{\sigma_X}$ que tenga media nula y varianza unidad, tal y como ya se ha visto para el caso discreto.

Además, podemos definir:

- Coeficiente de asimetría: $\gamma_1 = E\{\frac{(X-\eta_X)^3}{\sigma^3}\} = \frac{\mu_3}{\sigma_X^3}$
- Coeficiente de exceso de curtosis: $\gamma_2 = E\{\frac{(X-\eta_X)^4}{\sigma^4}\} - 3 = \frac{\mu_4}{\sigma_X^4} - 3$. La distribución normal tiene un exceso de curtosis de 0, razón por la que se resta 3.

Funciones condicionadas

Consideremos un espacio de probabilidad $\mathscr{E}$ en el que se ha definido una variable aleatoria X, su función de distribución $F_X(x)$ y su función de densidad $f_X(x)$. Considérese, además, un suceso $A \in \mathscr{F}$.

La **función de distribución de X condicionada por A** es:

$$F_X(x \;/\; A) = P(X \leq x \;/\; A) = \frac{P(\{X \leq x \bigcap A)\}}{P(A)}$$

Y la **función de densidad de X condicionada por A** es:

$$f_X(x \;/\; A) = \frac{dF_X(x \;/\; A)}{dx} = \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x/2 \;/\; A) - F_X(x - \Delta x/2 \;/\; A)}{\Delta x} = = \lim_{\Delta x \to 0} \frac{P\left(\{(x - \Delta x/2) < x \leq (x + \Delta\right.}{\Delta x}$$

En general el suceso condicionante A podemos definirlo con intervalos de la variable aleatoria. Suponiendo que $A = \{x_1 \leq X \leq x_2\}$:

$$F_X(x \;/\; \{x_1 \leq X \leq x_2\}) = \frac{P(\{X \leq x \bigcap \{x_1 \leq X \leq x_2\ )\}}{P(\{x_1 \leq X \leq x_2\})} =$$

$$= \begin{cases} 0 & x < x_1 \\ \frac{F_X(x) - F_X(x_1^-)}{F_X(x_2) - F_X(x_1^-)} & x_1 \leq x \leq x_2 \\ 1 & x > x_2 \end{cases}$$

$$f_X(x \,/\, \{x_1 \leq X \leq x_2\}) = \begin{cases} 0 & x < x_1 \\ \frac{f_X(x)}{F_X(x_2)-F_X(x_1^-)} & x_1 \leq x \leq x_2 \\ 0 & x > x_2 \end{cases}$$

Teorema de Bayes

Recordemos el **Teorema de Bayes** con dos sucesos:

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)}$$

Si $B = \{X \leq x\}$:

$$P(A \,/\, X \leq x) = \frac{F_X(x \,/\, A)P(A)}{F_X(x)} \iff F_X(x \,/\, A) = \frac{P(A \,/\, X \leq x)F_X(x)}{P(A)}$$

Si $B = \{x_1 < X \leq x_2\}$:

$$P(A \,/ \{x_1 < X \leq x_2\}) = \frac{F_X(x_2 \,/\, A) - F_X(x_1 \,/\, A)}{F_X(x_2) - F_X(x_1)} P(A)$$

Si $B = \{X = x\}$ y la variable aleatoria es continua nos encontramos con un problema para condicionar por B pues su probabilidad es nula. En tal caso consideramos que $B = \{x < X - \Delta x/2 \leq x + \Delta x/2\}$ y tomamos un paso al límite:

$$\begin{aligned} P(A \,/\, X = x) &= \lim_{\Delta x \to 0} P(A \,/ \{x - \Delta x/2 < X \leq x + \Delta x/2\}) \\ &= \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x/2 \,/\, A) - F_X(x - \Delta x/2 \,/\, A)}{F_X(x + \Delta x/2) - F_X(x - \Delta x/2)} P(A) \end{aligned}$$

Por tanto,

$$P(A \,/\, X = x) = \frac{f_X(x \,/\, A)P(A)}{f_X(x)} \iff f_X(x \,/\, A) = \frac{P(A \,/\, X = x)f_X(x)}{P(A)}$$

Teorema de la Probabilidad Total

Sea $A_i \in \mathscr{P}(\Omega)\; i = 1 \ldots K$ es una partición del espacio muestral Ω. Se obtiene inmediatamente el **Teorema de la Probabilidad Total**:

$$F_X(x) = F_X(x \,/\, A_1)P(A_1) + \ldots + F_X(x \,/\, A_K)P(A_K)$$

Derivando:

$$f_X(x) = f_X(x \,/\, A_1)P(A_1) + \ldots + f_X(x \,/\, A_K)P(A_K)$$

Podemos obtener una versión continua del teorema integrando $f_X(x \,/\, A) = \frac{P(A \,/\, X=x)f_X(x)}{P(A)}$:

$$P(A) = \int_{-\infty}^{+\infty} P(A \,/\, X = x)f_X(x)dx$$

Esperanza y momentos condicionados

$$\eta_{X/A} = E\{X \,/\, A\} = \int_{-\infty}^{\infty} x f_X(x \,/\, A) dx$$