### 3.6.2. Esperanza, varianza y desviación estándar de una variable borroso aleatoria

La primera definición que ofrecemos para la esperanza matemática, y las medidas de dispersión varianza y desviación estándar de una variable borroso aleatoria, es la propuesta en Kruse y Meyer (1987), y consiste en definirlas vía principio de extensión de Zadeh. De esta forma, estas medidas vendrán dadas también por los subconjuntos borrosos que notamos como $\tilde{E}[X]$, $\tilde{V}[X]$ y $\tilde{D}[X]$, cuyas funciones de pertenencia quedarán definidas a través de la de $\tilde{X}$. Así, obtenemos la función de pertenencia de la esperanza matemática como:

$$\mu_{\tilde{E}[X]}(x) = \bigvee_{x=E[X]} \mu_{\tilde{X}}(X)$$

la de la varianza será:

$$\mu_{\tilde{V}[X]}(x) = \bigvee_{x=V[X]} \mu_{\tilde{X}}(X)$$

y la de la desviación estándar:

$$\mu_{\tilde{D}[X]}(x) = \bigvee_{x=D[X]} \mu_{\tilde{X}}(X) = \mu_{\tilde{V}[X]}(x^2)$$

Como E[X], V[X] y D[X] hemos notado los operadores esperanza matemática, varianza y desviación estándar para una variable aleatoria cierta X.

Por otra parte, para una variable aleatoria discreta X cuyas realizaciones son: $\{x_1, x_2, ..., x_n\}$ la cuales presentan una probabilidad $\{p_1, p_2, ..., p_n\}$, podemos entender que E[X], V[X] y D[X] son funciones de sus realizaciones y las probabilidades que éstas llevan asociadas. Así, su expresión es:

$$E[X] = \sum_{i=1}^{n} p_i x_i$$

$$V[X] = \sum_{i=1}^{n} p_i \left( x_i - \sum_{i=1}^{n} p_i x_i \right)^2$$

$$D[X] = \sqrt{\sum_{i=1}^{n} p_i \left( x_i - \sum_{i=1}^{n} p_i x_i \right)^2}$$

Así, si entendemos como variable borroso aleatoria a una variable aleatoria cuya realización i-ésima es un número borroso $\tilde{A}_i$, siendo su probabilidad asociada $p_i$, i=1,2,...,n, la función de pertenencia de la esperanza, la varianza y la desviación estándar pueden ser halladas a través del principio de extensión de Zadeh realizando:

$$\mu_{\tilde{E}[X]}(x) = \bigvee_{x=\sum_{i=1}^{n} p_i x_i} \left(\mu_{\tilde{A}_1}(x_1) \wedge \mu_{\tilde{A}_2}(x_2) \wedge \ldots \wedge \mu_{\tilde{A}_n}(x_n)\right)$$

$$\mu_{\tilde{V}[X]}(x) = \bigvee_{x=\sum_{i=1}^{n} p_i \left(x_i - \sum_{i=1}^{n} p_i x_i\right)^2} \left(\mu_{\tilde{A}_1}(x_1) \wedge \mu_{\tilde{A}_2}(x_2) \wedge \ldots \wedge \mu_{\tilde{A}_n}(x_n)\right)$$

$$\mu_{\tilde{D}[X]}(x) = \bigvee_{x=\sqrt{\sum_{i=1}^{n} p_i \left(x_i - \sum_{i=1}^{n} p_i x_i\right)^2}} \left(\mu_{\tilde{A}_1}(x_1) \wedge \mu_{\tilde{A}_2}(x_2) \wedge \ldots \wedge \mu_{\tilde{A}_n}(x_n)\right) = \mu_{\tilde{V}[X]}(x^2)$$

Es evidente que en muchas ocasiones no podremos hallar la expresión analítica de la función de pertenencia de $\tilde{E}[X]$, $\tilde{V}[X]$ y $\tilde{D}[X]$, pero sí que será posible representar dichos momentos de una variable borroso aleatoria a través de sus $\alpha$-cortes. Así, los $\alpha$-cortes de la esperanza matemática, $E[X]_\alpha$ vienen dados por:

$$E[X]_\alpha = \left\{ x \mid x = \sum_{i=1}^{n} p_i x_i ,\, x_i \in A_{i\alpha} ,\, i = 1,2,\ldots,n \right\}$$

Y como la función esperanza matemática es monótona creciente respecto al valor de las realizaciones, obtenemos que:

$$E[X]_\alpha = \left[E^1[X](\alpha), E^2[X](\alpha)\right] = \left[E\left[X^1(\alpha)\right], E\left[X^2(\alpha)\right]\right] = \left[\sum_{i=1}^{n} p_i A_i^1(\alpha), \sum_{r=1}^{n} p_i A_i^2(\alpha)\right]$$

Obsérvese que su valor coincide con el que se obtiene para el valor medio de un haz de números borrosos.

Respecto a la varianza, observamos que sus $\alpha$-cortes son:

$$V[Z]_\alpha = \left\{ x \in R^+ \mid x = \sum_{i=1}^{n} p_i \left( x_i - \sum_{i=1}^{n} p_i x_i \right)^2 ,\, x_i \in A_{i_\alpha} \; i = 1,2,\ldots,n \right\} =$$

$$=[\text{Min } x, \text{Max } x]$$

Dado que no existe monotonía de V[X] respecto a las realizaciones, x, para hallar los extremos de $V[X]_\alpha$ deberemos plantear el siguiente programa matemático:

Minimizar (Maximizar) $x = \sum_{i=1}^{n} p_i \left( x_i - \sum_{i=1}^{n} p_i x_i \right)^2$

s.a.:

$$A_i^1(\alpha) \le x_i \le A_i^2(\alpha) \quad i=1,..,n$$

Siendo el máximo de x, el extremo superior de $V[X]_\alpha$, y el mínimo de x, su extremo inferior. Si bien, estos programas matemáticos son relativamente sencillos de resolver, realizar dicha implementación para todos los que se desee, puede ser costoso. Por ejemplo, si se trabaja con una escala endecadaria, el número de programas matemáticos a resolver será 22. Inspirándonos en Frühwirth-Schnatter (1992), donde se propone una aproximación a la desviación estándar muestral cuando ésta se compone de números borrosos basada en la aritmética de los intervalos de confianza, apuntaremos una aproximación a los $\alpha$-cortes de la varianza que facilitarán su computación, aunque a costa de obtener un número borroso más incierto.

En certeza, podemos definir a una realización centrada de una variable aleatoria X como: $c_i=x_i-E[X]$, siendo entonces la variable C, C=X- E[X], la variable aleatoria centrada que deriva de X. De forma análoga, definimos para la i-ésima realización de $\tilde{X}$, $\tilde{A}_i$, por tanto, la correspondiente a una variable borroso aleatoria, a su valor centrado como $\tilde{C}_i = \tilde{A}_i - \tilde{E}[X]$, utilizándose en la construcción de los $\alpha$-cortes de $\tilde{C}_i$, $C_{i\alpha}$ la aritmética de intervalos de confianza se tiene:

$$C_{i_\alpha} = \left[C_i^1(\alpha), C_i^2(\alpha)\right] = \left[A_i^1(\alpha) - E^2[X](\alpha), A_i^2(\alpha) - E^1[X](\alpha)\right]$$

A partir de los $\alpha$-cortes de $\tilde{C}_i$, hallamos $\left(\tilde{C}_i\right)^2$, que es un número borroso cuyos $\alpha$-cortes son:

$$\left(C_{i_\alpha}\right)^2 = \left[\left[(C_i)^2\right]^-(\alpha), \left[(C_i)^2\right]^+(\alpha)\right]$$

donde:

$$\left[(C_i)^2\right]^-(\alpha) = \begin{cases} \left[C_i^1(\alpha)\right]^2 & \text{si } C_i^1(\alpha) \ge 0 \\ \left[C_i^2(\alpha)\right]^2 & \text{si } C_i^2(\alpha) \le 0 \\ 0 & \text{en otro caso} \end{cases} \quad \text{y} \quad \left[(C_i)^2\right]^+(\alpha) = \begin{cases} \left[C_i^2(\alpha)\right]^2 & \text{si } C_i^1(\alpha) + C_i^2(\alpha) \ge 0 \\ \left[C_i^1(\alpha)\right]^2 & \text{en otro caso} \end{cases}$$

De esta forma,

$$V[X]_\alpha \approx \sum_{i=1}^n p_i \left(C_{i_\alpha}\right)^2 = \left[\sum_{i=1}^n p_i \left[(C_i)^2\right]^-(\alpha), \sum_{i=1}^n p_i \left[(C_i)^2\right]^+(\alpha)\right]$$

y así, como $D[X]_\alpha = \sqrt{V[X]_\alpha}$, entonces:

$$D[X]_\alpha \approx \left[\sqrt{\sum_{i=1}^{n} p_i \left[(C_i)^2\right]^-(\alpha)}, \sqrt{\sum_{i=1}^{n} p_i \left[(C_i)^2\right]^+(\alpha)}\right]$$

El concepto de varianza y desviación estándar que hemos planteado anteriormente está basado en el concepto de distancia euclídea borrosa entre números borrosos –ver para el análisis de dicho concepto Dubois y Prade (1980)-. Sin embargo, diversos autores como Feng *et al.* (2000) consideran que, aunque es razonable sostener que el valor esperado de una variable borroso aleatoria es un número borroso, las varianza de una variable borroso aleatoria no debe ser borrosa, ya que por la naturaleza del indicador varianza o desviación estándar éstos deben ser índices ciertos. Así, la idea subyacente de esta modelización de varianza es la utilización del concepto de distancia cierta entre números borrosos, concepto que ya propusimos en el apartado 1.4. como instrumento en la aproximación de números borrosos. Basándose en el concepto de distancia euclídea cierta entre números borrosos, Feng *et al*. definen como varianza cierta de una variable borroso aleatoria, la cual notaremos como $V^*[X]$ a:

$$V^*[X] = \frac{1}{2}\int_0^1 \left(V\left[X^1(\alpha)\right] + V\left[X^2(\alpha)\right]\right) d\alpha =$$

$$= \frac{1}{2}\int_0^1 \left\{\sum_{i=1}^{n} p_i \left(A_i^1(\alpha) - \sum_{i=1}^{n} p_i A_i^1(\alpha)\right)^2 + \sum_{i=1}^{n} p_i \left(A_i^2(\alpha) - \sum_{i=1}^{n} p_i A_i^2(\alpha)\right)^2\right\} d\alpha$$

Siendo entonces la desviación estándar que se deriva del concepto de varianza de Feng, $D^*[X]$:

$$D^*[X] = \sqrt{V^*[X]} = \sqrt{\frac{1}{2}\int_0^1 \left(V\left[X^1(\alpha)\right] + V\left[X^2(\alpha)\right]\right) d\alpha}$$

### 3.6.3. Función de distribución y cuantiles de una variable borroso aleatoria

Para una variable aleatoria discreta X, cuyas realizaciones son $\{x_1, x_2,\ldots,x_n\}$ y cuyas probabilidades asociadas son $\{p_1, p_2,\ldots,p_n\}$, la función de distribución de X, que notamos como F[x] es una función tal que:

$$F[x] = p[X \le x] = \sum_{\forall i | x_i \le x} p_i$$

Una función de distribución F[x] debe cumplir las siguientes propiedades:

i) F[x] es no decreciente

ii) $F[x] \to 0$ si $x \to -\infty$ y $F[x] \to 1$ si $x \to 0$

iii) $p[X>x] = 1-p[X\le x] = 1- F[x]$

En el caso de variables borroso aleatorias, la función de distribución será un subconjunto borroso $\tilde{F}[x]$, que indicará p[ $\tilde{X} \leq x$ ]. Su función de pertenencia se podrá hallar a través de la de $\tilde{X}$ planteando:

$$\mu_{\tilde{F}[x]}(z) = \bigvee_{z=F[x]=p(X\leq x)} \mu_{\tilde{X}}(X)$$

Asimismo, también podemos hallar sus conjuntos de nivel, lo cual será normalmente más operativo. En este caso, debemos tener en cuenta que dos variables aleatorias X e Y discretas, cuyo campo de variación es, $\{x_i\}_{i=1\ldots n}$ y $\{y_i\}_{i=1\ldots n}$ respectivamente con $P[X=x_i]=P[Y=y_i]=p_i$, si $x_i \geq y_i$ $\forall i$, se cumple siempre que $P[X \leq x] \leq P[Y \leq x]$. De esta forma, para $X_\alpha$ definiremos los α-cortes de la función de distribución, $F[x]_\alpha$ a través de los α-cortes de la variable aleatoria inferior y superior para dicho nivel α. En concreto, el extremo inferior $F^1[X](\alpha)$, se construirá a través de $X^2(\alpha)$ y el superior $F^2[X](\alpha)$ a través de $X^1(\alpha)$:

$$F^2[x](\alpha) = p[X^1(\alpha) \leq x] = \sum_{\forall i / A_i^1(\alpha) \leq x} p_i \text{ y } F^1[x](\alpha) = p[X^2(\alpha) \leq x] = \sum_{\forall i / A_2^2(\alpha) \leq x} p_i$$

Así, $F[x]_\alpha$ no será, en variables borroso aleatorias discretas, un intervalo de confianza, ya que su dominio es discreto. Sin embargo se puede afirmar que:

$$F[x]_\alpha \subseteq [F^1[x](\alpha), F^2[x](\alpha)]$$

cumpliéndose asimismo que $F[x]_\alpha \supseteq F[x]_{\alpha'}\ \forall \alpha \leq \alpha'$ ya que $A_{i\alpha} \supseteq A_{i\alpha'}$, i=1,2,...,n.

Por supuesto, también podremos hallar la función de pertenencia de $\tilde{F}[x]$ través de $F[x]_\alpha$ como:

$$\mu_{\tilde{F}[x]}(z) = \{\text{Max } \alpha \mid z \in F[X]_\alpha\} \text{, con } \alpha \in [0,1]$$

Para definir el concepto de cuantil en una variable borroso aleatoria también partiremos del concepto de cuantil para una variable aleatoria discreta. Se denomina ε-cuantil de una variable aleatoria X al menor valor que acumula a su izquierda una probabilidad de ε, siendo dicho cuantil notado como $q^\varepsilon$. En concreto, éste se halla como:

$$q^\varepsilon = \text{Min } x|F[x]\geq\varepsilon \text{ o } \text{Min } x \mid p[X \leq x] \geq \varepsilon$$

Por supuesto, en una variable borroso aleatoria los cuantiles serán subconjuntos borrosos. A partir de la función de pertenencia de $\tilde{X}$, podemos definir el ε-cuantil de $\tilde{X}$, $\tilde{Q}^\varepsilon$ como:

$$\mu_{\tilde{Q}^{\varepsilon}}(q) = \bigvee_{q=\text{Min } x|F[x]\geq\varepsilon} \mu_{\tilde{X}}(X)$$

A pesar de que los $\alpha$-cortes de la función de distribución no vienen estrictamente dados mediante un intervalo de confianza, los $\alpha$-cortes del 1-$\varepsilon$ cuantil de $\tilde{Q}^{\varepsilon}$, $Q^{\varepsilon}_{\alpha}$, si lo serán. De esta forma para hallar $Q^{\varepsilon}_{\alpha}$ deberemos partir de $F[x]_{\alpha}$, obteniéndose:

$$Q^{\varepsilon}_{\alpha} = \left[Q^{1\varepsilon}(\alpha), Q^{2\varepsilon}(\alpha)\right]$$

donde:

$$Q^{1\varepsilon}(\alpha) = \text{Min } x \mid F^{2}[x](\alpha) \geq \varepsilon \quad \text{y} \quad Q^{2\varepsilon}(\alpha) = \text{Min } x \mid F^{1}[x](\alpha) \geq \varepsilon$$

Es decir, a partir de la inversa de la función de distribución inferior, que es la de $X^{2}(\alpha)$, obtendremos el extremo superior del $\alpha$-corte de $\tilde{Q}^{\varepsilon}$ y viceversa.

Asimismo, si las realizaciones de las variables borroso aleatorias $\tilde{A}_1$, $\tilde{A}_2$,…, $\tilde{A}_n$ cumplen que:

$A^{1}_{i}(\alpha) \leq A^{1}_{i+1}(\alpha)$, i=1,2,…,n-1 $\forall\alpha$

$A^{2}_{i}(\alpha) \leq A^{2}_{i+1}(\alpha)$, i=1,2,…,n-1 $\forall\alpha$

Los extremos inferior y superior de $Q^{\varepsilon}_{\alpha}$ son:

$$Q^{1,\varepsilon}(\alpha) = A^{1}_{i}(\alpha) \mid \underset{i}{\text{Min}} \sum_{j\leq i} p_j \geq \varepsilon \ \text{ y } \ Q^{2,\varepsilon}(\alpha) = A^{2}_{i}(\alpha) \mid \underset{i}{\text{Min}} \sum_{j\leq i} p_j \geq \varepsilon$$

Y así,

$$Q^{\varepsilon}_{\alpha} = \left[Q^{1\varepsilon}(\alpha), Q^{2\varepsilon}(\alpha)\right] = \left[A^{1}_{i}(\alpha), A^{1}_{i}(\alpha)\right]$$

y por tanto:

$$\tilde{Q}^{\varepsilon} = \tilde{A}_i$$

### 3.6.4. Inferencia estadística en muestras donde las observaciones vienen dadas por números borrosos

En este epígrafe analizaremos la determinación del la media muestral, de la varianza y desviación estándar muestral, la función de distribución empírica y los cuantiles empíricos en muestras tomadas de variables aleatorias, cuando las cuantías que son observadas no vienen dadas de forma cierta, sino mediante números borrosos. Es decir, en este caso, se supone que el observador dispone de n observaciones que vienen dadas por números borrosos, siendo el correspondiente a la i-ésima observación $\widetilde{A}_i$ :

$$\widetilde{A}_i = \{x, \mu_{\widetilde{A}_i}(x)\} = \{A_{i_\alpha} = [A_i^1(\alpha), A_i^2(\alpha)] \mid 0 \le \alpha \le 1\},\ i=1,2,...,n.$$

El hecho de que los datos obtenidos de la muestra sean borrosos, puede venir dado, por ejemplo, por la dificultad o la imposibilidad de cuantificar exactamente los mismos. Ello puede darse, por ejemplo, si la metodología que se utiliza para su medición no es la adecuada o bien por la misma naturaleza de los datos –por ejemplo, en los mercados financieros, durante una misma sesión no suele negociarse un único precio, sino que suelen negociarse varios, de forma que, por ejemplo, el precio de una acción un determinado día no es “2000”, sino “aproximadamente 2000”-. En estas circunstáncias la cuantificación de dichos datos es vaga, siendo más realista modelizarlos a través de números borrosos.

En nuestra exposición, nos basaremos en el trabajo ya mencionado de Frühwirth-Schnatter (1992). El enfoque que en dicho trabajo se da al problema que se analiza, es muy similar al expuesto por nosotros en el análisis de las variables borroso aleatorias, siendo esta la razón por la que incluimos esta herramienta dentro del apartado dedicado a las variables borroso aleatorias.

### *3.6.4.1. Media muestral, varianza muestral y desviación estándar muestral*

Estos parámetros representativos del valor medio y la dispersión serán definidos a través del principio de extensión de Zadeh. Para ello partiremos de que la esperanza muestral de una variable aleatoria X, $\hat{E}[X]$, su varianza, $\hat{V}[X]$ y su desviación estándar, $\hat{D}[X]$, si la muestra obtenida es un vector $(x_1, x_2,...,x_n)$, siendo sus componentes cuantías ciertas, puede expresarse como una función $f:R^n \rightarrow R$, en la que las variables independientes son los datos obtenidos. En concreto:

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$V[X] = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$$

$$D[X] = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2}$$

Sin embargo, si los datos vienen dados a través de un vector borroso, donde el i-ésimo componente es el número borroso $\tilde{A}_i$, i=1,2,…,n, la esperanza, la varianza y la desviación estándar muestral son unos números borrosos que notamos respectivamente como $\tilde{\hat{E}}[X]$, $\tilde{\hat{V}}[X]$ y $\tilde{\hat{D}}[X]$. Su función de pertenencia puede ser hallada vía principio de extensión de Zadeh como:

$$\mu_{\hat{\tilde{E}}[X]}(x) = \bigvee_{x=\frac{1}{n}\sum_{i=1}^{n} x_i} \left(\mu_{\tilde{A}_1}(x_1) \wedge \mu_{\tilde{A}_2}(x_2) \wedge ... \wedge \mu_{\tilde{A}_n}(x_n)\right)$$

$$\mu_{\hat{\tilde{V}}[X]}(x) = \bigvee_{x=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2} \left(\mu_{\tilde{A}_1}(x_1) \wedge \mu_{\tilde{A}_2}(x_2) \wedge ... \wedge \mu_{\tilde{A}_n}(x_n)\right)$$

$$\mu_{\hat{\tilde{D}}[X]}(x) = \bigvee_{x=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2}} \left(\mu_{\tilde{A}_1}(x_1) \wedge \mu_{\tilde{A}_2}(x_2) \wedge ... \wedge \mu_{\tilde{A}_n}(x_n)\right) = \mu_{\hat{\tilde{D}}[X]}(x^2)$$

Por otra parte, los α-cortes de la esperanza matemática de $\tilde{\hat{E}}[X]$, $\hat{E}[X]_\alpha$, son intervalos de confianza que se pueden hallar como:

$$\hat{E}[X]_\alpha = \left\{ x \mid x = \frac{1}{n}\sum_{i=1}^{n} x_i, x_i \in A_{i\alpha}, i = 1,2,...,n \right\}$$

Y como la media muestral es monótona creciente respecto al valor de las cuantías que la forman obtenemos que:

$$\hat{E}[X]_\alpha = \left[\hat{E}^1[X](\alpha), \hat{E}^2[X](\alpha)\right] = \left[\frac{1}{n}\sum_{i=1}^{n} A_i^1(\alpha), \frac{1}{n}\sum_{r=1}^{n} A_i^2(\alpha)\right]$$

Respecto a la varianza muestral, observamos que:

$$\hat{V}[X]_\alpha = \left\{ x \in R^+ \mid x = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2, x_i \in A_{i_\alpha}\ i = 1,2,...,n \right\} =$$

$$= [\text{Min } x, \text{Max } x]$$

Dado que no existe monotonía de $\hat{V}[X]$ respecto al valor de los valores obtenidos en la muestra, para hallar los extremos de $\hat{V}[X]_\alpha$ deberemos plantear y resolver el siguiente programa matemático:

Minimizar (Maximizar) $x = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \dfrac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

s.a.:

$A_i^1(\alpha) \leq x_i \leq A_i^2(\alpha)$ i=1,..,n

Siendo el máximo de x, el extremo superior de $\hat{V}[X]_\alpha$, y el mínimo de x, su extremo inferior. Como ya comentamos cuando analizamos la determinación de los α-cortes de la varianza borrosa de una variable borroso aleatoria, creemos que, si la escala de verdad que emplea el observador es relativamente detallada, este enfoque puede no ser operativo, ya que el número de programas a resolver puede ser excesivo. Frühwirth-Schnatter (1992) propone una aproximación a la varianza y a la desviación estándar muestral, que se basa en la aritmética de intervalos, siendo dicha aproximación la que inspiró nuestra propuesta para aproximar la varianza de una variable borroso aleatoria. Esta aproximación incluye al α-corte real de la varianza, y facilita su cálculo, si bien a costa de obtenerse un número borroso más incierto.

Para ello, deberemos hallar el valor centrado de cada una de las observaciones de forma análoga a lo realizado en 3.6.2. Para la i-ésima realización de la muestra $\tilde{A}_i$, su valor centrado es $\tilde{C}_i = \tilde{A}_i - \hat{\tilde{E}}[X]$. Para hallar los α-cortes de $\tilde{C}_i$ utilizamos la aritmética de intervalos, de forma que:

$$C_{i_\alpha} = \left[C_i^1(\alpha), C_i^2(\alpha)\right] = \left[A_i^1(\alpha) - \hat{E}^2[X](\alpha), A_i^2(\alpha) - \hat{E}^1[X](\alpha)\right]$$

A partir de los α-cortes de $\tilde{C}_i$, hallamos $\left(\tilde{C}_i\right)^2$, siendo los α-cortes de este último número borroso:

$$\left(C_{i_\alpha}\right)^2 = \left[\left[(C_i)^2\right]^-(\alpha), \left[(C_i)^2\right]^+(\alpha)\right]$$

donde:

$$\left[(C_i)^2\right]^-(\alpha) = \begin{cases} \left[C_i^1(\alpha)\right]^2 & \text{si } C_i^1(\alpha) \geq 0 \\ \left[C_i^2(\alpha)\right]^2 & \text{si } C_i^2(\alpha) \leq 0 \\ 0 & \text{en otro caso} \end{cases} \quad \text{y} \quad \left[(C_i)^2\right]^+(\alpha) = \begin{cases} \left[C_i^2(\alpha)\right]^2 & \text{si } C_i^1(\alpha) + C_i^2(\alpha) \geq 0 \\ \left[C_i^1(\alpha)\right]^2 & \text{en otro caso} \end{cases}$$

De esta forma,

$$\hat{V}[X]_\alpha \approx \frac{1}{n-1}\sum_{i=1}^{n}\left(C_{i_\alpha}\right)^2 = \left[\frac{1}{n-1}\sum_{i=1}^{n}\left[(C_i)^2\right]^-(\alpha), \frac{1}{n-1}\sum_{i=1}^{n}\left[(C_i)^2\right]^+(\alpha)\right]$$

y así, como $D[X]_\alpha = \sqrt{V[X]_\alpha}$, entonces:

$$D[X]_\alpha \approx \left[\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left[(C_i)^2\right]^-(\alpha)}, \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left[(C_i)^2\right]^+(\alpha)}\right]$$

*3.6.4.2. Función de distribución muestral y cuantiles muestrales*

Para definir la función de distribución y los cuantiles para muestras con observaciones borrosas, partiremos de su definición suponiendo que la muestra tomada se compone de cuantías ciertas, las cuales vienen dadas por el vector $(x_1,x_2,...,x_i,...,x_n)$. El valor de la función de distribución empírica para una cuantía x, $\hat{F}[x]$ es:

$$\hat{F}[x] = \sum_{\forall x_i \leq x} \frac{1}{n}$$

siendo el dominio de $\hat{F}[x]$ el conjunto discreto $\left\{0,\frac{1}{n},\frac{2}{n},...,1\right\}$.

Si la muestra viene dada por n valores borrosos $\tilde{A}_i$, i=1,2,...,n, entonces, la función de distribución para un valor x viene dado por un subconjunto borroso $\tilde{\hat{F}}[x]$ cuya función de pertenencia es:

$$\mu_{\tilde{\hat{F}}[x]}(z) = \bigvee_{z=\sum_{\forall x_i \leq x} \frac{1}{n}} \left(\bigwedge_{i=1}^{n} \mu_{\tilde{A}_i}(x_i)\right)$$

También podemos hallar los α-cortes de la función de distribución muestral a través de sus α-cortes, $\hat{F}[x]_\alpha$. El valor inferior de $\hat{F}[x]_\alpha$, $\hat{F}^1[x](\alpha)$ se hallará a través de los extremos superiores de los α-cortes de las observaciones, $A_i^2(\alpha)$, i=1,2,...,n y el superior, $\hat{F}^2[x](\alpha)$, a través de los inferiores, $A_i^1(\alpha)$, i=1,2,...,n. Así:

$$\hat{F}^2[x](\alpha) = \sum_{\forall A_i^1(\alpha) \leq x} \frac{1}{n} \text{ y } \hat{F}^1[x](\alpha) = \sum_{\forall A_i^2(\alpha) \leq x} \frac{1}{n}$$

Así, $\hat{F}[x]_\alpha$ no será un intervalo de confianza, ya que como se ha comentado, su dominio es discreto, $\left\{0,\frac{1}{n},\frac{2}{n},...,1\right\}$ pero si se cumple que:

$$\hat{F}[x]_\alpha \subseteq \left[\hat{F}^1[x](\alpha), \hat{F}^2[x](\alpha)\right]$$

cumpliéndose asimismo que $\hat{F}[x]_\alpha \supseteq \hat{F}[x]_{\alpha'}$ $\forall \alpha \leq \alpha'$ ya que $A_{i\alpha} \supseteq A_{i\alpha'}$, i=1,2,...,n

Por supuesto, podremos hallar la función de pertenencia de $\hat{\tilde{F}}[x]$ través de $\hat{F}[x]_\alpha$ como:

$$\mu_{\tilde{F}[x]}(z) = \{\text{Max}\ \alpha \mid z \in \hat{F}[X]_\alpha\}\text{, con } \alpha \in [0,1]$$

Por otra parte, para definir el concepto de cuantil borroso en una muestra donde las observaciones son borrosas, partiremos del concepto de cuantil en el caso en que la muestra sea un vector cierto $(x_1, x_2,...,x_i,...,x_n)$ donde $x_1 \leq x_2 \leq ... \leq x_r \leq ... \leq x_n$

El cuantil de orden ε, $\hat{q}^\varepsilon$ es aquel valor $x_{i+1}$ que presenta una frecuencia acumulada tal que:

$$\hat{q}^\varepsilon = x_{i+1} \mid \hat{F}[x_i] = \frac{i}{n} < \varepsilon \leq \hat{F}[x_{i+1}] = \frac{i+1}{n}$$

Si las observaciones de la muestra son borrosas, $\tilde{A}_i$, i=1,2,...,n, entonces, dicho cuantil será también un número borroso, que notamos como $\hat{\tilde{Q}}^\varepsilon$, donde la posibilidad de un valor del mismo $q^\varepsilon = x_{i+1}$ vendrá dado por:

$$\mu_{\hat{\tilde{Q}}^\varepsilon}(q) = \bigvee_{q=x_{i+1} \mid F[x_i] < \varepsilon \leq F[x_{i+1}]} \left( \bigwedge_{i=1}^{n} \mu_{\tilde{A}_i}(x_i) \right)$$

Como en el caso de las variables borroso aleatorias, los α-cortes del cuantil de orden ε pueden ser hallados a través de los α-cortes de las observaciones que componen la muestra estudiada. De esta forma, para hallar $\hat{Q}^\varepsilon_\alpha$, donde:

$$\hat{Q}^\varepsilon_\alpha = \left[\hat{Q}^{1\varepsilon}(\alpha), \hat{Q}^{2\varepsilon}(\alpha)\right]$$

deberemos hallar el valor inferior de su α-corte $\hat{Q}^{1\varepsilon}(\alpha)$ a través de los extremos inferiores de los α-cortes con dicho nivel de presunción de $\tilde{A}_i$, i=1,2,...,n, $A_i^1(\alpha)$, y el superior, $\hat{Q}^{2\varepsilon}(\alpha)$, a través de los superiores, $A_i^2(\alpha)$, i=1,2,...,n. Para ello ordenamos los extremos de los α-cortes de $\tilde{A}_1$, $\tilde{A}_2, \ldots, \tilde{A}_n$ para cada nivel en orden creciente, de forma que:

$$A_r^1(\alpha) \leq A_{r+1}^1(\alpha)\text{, i=0,1,…,n-1 } \forall\alpha$$

$$A_r^2(\alpha) \leq A_{r+1}^2(\alpha)\text{, i=0,1,…,n-1 } \forall\alpha$$

donde tras la ordenación efectuada, el i-ésimo extremo inferior que se obtiene no tiene por que pertenecer al mismo número borroso que el i-ésimo extremo superior.

Podemos finalmente obtener los extremos del $\alpha$-corte del cuantil que pretendemos hallar de forma que $\hat{Q}^{1,\varepsilon}(\alpha)^{\varepsilon}$ es aquel extremo $A^{1}_{i+1}(\alpha)$ y $\hat{Q}^{2,\varepsilon}(\alpha)$ es aquel extremo $A^{2}_{i+1}(\alpha)$ que cumplen respectivamente:

$$\hat{F}^{2}\left[A^{1}_{i}(\alpha)\right]=\frac{i}{n}<\varepsilon\leq\hat{F}^{2}\left[A^{1}_{i+1}(\alpha)\right]=\frac{i+1}{n} \text{ y } \hat{F}^{1}\left[A^{2}_{i}(\alpha)\right]=\frac{i}{n}<\varepsilon\leq\hat{F}^{1}\left[A^{2}_{i+1}(\alpha)\right]=\frac{i+1}{n}$$

## 3.7. VARIABLES ALEATORIAS NORMALES CON MEDIA Y VARIANZA BORROSAS

El concepto de variable normal borrosa ha sido utilizado, entre otros por Yager (1982) para realizar predicciones con una regresión lineal convencional cuando las variables independientes vienen dadas por números borrosos. En este caso, aunque el término de error asociado a la regresión es una variable aleatoria normal de media 0 y varianza constante, el valor de la predicción será una variable aleatoria normal, pero en este caso con media y varianza borrosa, ya que el valor de estos parámetros depende del valor de la media y varianza de la predicción, que vienen dados por cuantías difusas. Asimismo, en Wierzchon (1988) se utiliza dicho concepto en un contexto de programación lineal borroso aleatoria.

Por otra parte, también podríamos encontrarnos ante una cuantía aleatoria normal y borrosa a la vez cuando, por ejemplo, si conocemos que un determinado fenómeno se comporta según una variable aleatoria normal, pero su media y desviación estándar debe ser inferida a través de una muestra donde las observaciones son difusas. La estimación en este caso de la media y la desviación estándar muestral ha sido analizada en el apartado 3.6.4.1. Empecemos por definir en un ambiente de certeza las cuestiones que después analizaremos en borrosidad.

### 3.7.1. Función de distribución, cuantiles e intervalos de confianza para variables aleatorias normales

Para una variable aleatoria X, que se comporta según una normal de media m y desviación estándar $\sigma$, $X\sim N(m,\sigma)$, podemos hallar el valor de distribución de un valor x, $F[x] = p[X\leq x]$, tipificando dicho valor:

$$z=\frac{x-m}{\sigma}$$

y como

$$Z=\frac{X-m}{\sigma}\sim N(0,1)$$

hallamos el valor de F[x] a través de la función de distribución de la normal tipificada sin más que:

$$p[X \leq x] = p[Z \leq z] = \Phi\left(\frac{x-m}{\sigma}\right)$$

asimismo como $\Phi(\cdot)$ notamos a la función de probabilidad acumulada de una variable aleatoria normal tipificada.

Por otra parte, dado que el cuantil de orden $\varepsilon$ de X, $q^{\varepsilon}$, es aquél valor que acumula a su izquierda una probabilidad de $\varepsilon$, es decir: $F[q^{\varepsilon}] = \varepsilon$, o de forma equivalente, $q^{\varepsilon}$ es un valor que cumple:

$$\Phi\left(\frac{q^{\varepsilon}-m}{\sigma}\right) = \varepsilon$$

podemos hallar el $\varepsilon$ cuantil de X a través de la función inversa de la normal tipificada como:

$$q^{\varepsilon} = m + \sigma\Phi^{-1}(\varepsilon)$$

Por último, para una variable aleatoria normal, un intervalo de confianza con una probabilidad de ocurrencia $\delta$, $I^{\delta}$ viene dado por aquéllos valores $x \in [a, b]$ que cumplen

$$P[a \leq x \leq b] = \delta$$

y:

$$a = m - \sigma\Phi^{-1}(\varepsilon) \text{ y } b = m + \sigma\Phi^{-1}(\varepsilon),$$

siendo $\varepsilon = 1 - \frac{1-\delta}{2}$

### 3.7.2. Función de distribución, cuantiles e intervalos de confianza para variables aleatorias normales con media y desviación estándar borrosas

En el problema que analizamos a continuación, tanto la media y la desviación estándar de la variable aleatoria que analizamos no vienen dadas de forma nítida, (por ejemplo, m=10 y $\sigma$ =5), sino que vienen cuantificadas mediante números borrosos (por ejemplo, m es aproximadamente 10 y $\sigma$ es aproximadamente 5). De esta forma, el instrumento de que disponemos para hacer predicciones es una variable aleatoria normal con media y desviación estándar borrosas que notaremos como $\tilde{X}$~N($\tilde{m},\tilde{\sigma}$). Al venir dadas tanto la media como la desviación estándar por números borrosos, podemos observar que:

$$\tilde{m} = \{x, \mu_{\tilde{m}}(x)\} = \{m_\alpha = [m^1(\alpha), m^2(\alpha)] \mid 0 \le \alpha \le 1\}$$

y

$$\tilde{\sigma} = \{x, \mu_{\tilde{\sigma}}(x)\} = \{\sigma_\alpha = [\sigma^1(\alpha), \sigma^2(\alpha)] \mid 0 \le \alpha \le 1\}$$

Así pues, N($\tilde{m},\tilde{\sigma}$) es un subconjunto borroso cuyo referencial es el de las variables aleatorias normales cuya función de pertenencia es:

$$\mu_{N(\tilde{\mu},\tilde{\sigma})}(y) = \bigvee_{y=N(x_1,x_2)} \left(\mu_{\tilde{m}}(x_1) \wedge \mu_{\tilde{\sigma}}(x_2)\right)$$

donde y es una variable aleatoria normal.

Para hallar la probabilidad que acumula un valor x cierto, $\tilde{F}[x]$, hallaremos en primer lugar su valor tipificado, en este caso será un número borroso $\tilde{z}$. La función de pertenencia de dicho valor $\tilde{z}$ será pues:

$$\mu_{\tilde{z}}(u) = \bigvee_{u=\frac{x-x_1}{x_2}} \left(\mu_{\tilde{m}}(x_1) \wedge \mu_{\tilde{\sigma}}(x_2)\right)$$

y sus $\alpha$-cortes:

$$z_\alpha = [z^1(\alpha), z^2(\alpha)]$$

donde:

$$z^1(\alpha) = \begin{cases} \dfrac{x - m^2(\alpha)}{\sigma^2(\alpha)} & \text{si } x - m^2(\alpha) \ge 0 \\ \dfrac{x - m^2(\alpha)}{\sigma^1(\alpha)} & \text{si } x - m^2(\alpha) < 0 \end{cases} \quad \text{y} \quad z^2(\alpha) = \begin{cases} \dfrac{x - m^1(\alpha)}{\sigma^1(\alpha)} & \text{si } x - m^1(\alpha) \ge 0 \\ \dfrac{x - m^1(\alpha)}{\sigma^2(\alpha)} & \text{si } x - m^1(\alpha) < 0 \end{cases}$$

Así, la función de distribución $\tilde{F}[x]$ que indicará la probabilidad P[$\tilde{X} \le x$], puede hallarse a través de la función de pertenencia del valor tipificado, $\tilde{z}$ como:

$$\mu_{\tilde{F}[x]}(v) = \bigvee_{v=\Phi(v)} \mu_{\tilde{z}}(u) = \bigvee_{v=\Phi\left(\frac{x-x_1}{x_2}\right)} \left(\mu_{\tilde{m}}(x_1) \wedge \mu_{\tilde{\sigma}}(x_2)\right)$$

Asimismo, sus $\alpha$-cortes, $F[x]_\alpha$ se hallan fácilmente, ya que $\Phi(\cdot)$ es monótona creciente respecto de su argumento. De esta forma:

$$F[x]_\alpha = \left[F[x]^1(\alpha), F[x]^2(\alpha)\right] = \left[\Phi^{-1}\left[z^1(\alpha)\right], \Phi^{-1}\left[z^2(\alpha)\right]\right]$$

Por otra parte, a través del principio de extensión de Zadeh podemos definir, el cuantil de orden ε de $\tilde{X}$, que notaremos como $\tilde{Q}^\varepsilon$. Su función de pertenencia será:

$$\mu_{\tilde{Q}^\varepsilon}(q) = \bigvee_{q = x_1 + x_2\Phi^{-1}(\varepsilon)} \left(\mu_{\tilde{m}}(x_1) \wedge \mu_{\tilde{\sigma}}(x_2)\right)$$

Y dado que la función q que evaluamos (el cuantil) es creciente respecto a la media y la varianza de la variable aleatoria, sus α-cortes son:

$$Q_\alpha^\varepsilon = \left[m^1(\alpha) - \sigma^1(\alpha)\Phi^{-1}(\varepsilon), m^2(\alpha) + \sigma^2(\alpha)\Phi^{-1}(\varepsilon)\right]$$

Por último, el intervalo de confianza con una probabilidad de ocurrencia δ, será un subconjunto borroso cuyo referencial son los intervalos de confianza, por tanto, no es un número borroso, y que notaremos como $\tilde{I}^\delta$. Su función de pertenencia se halla, como siempre, aplicando el principio de extensión, de forma que:

$$\mu_{\tilde{I}^\delta}(I) = \bigvee_{I = \left[x_1 - x_2\Phi^{-1}(\varepsilon), x_1 + x_2\Phi^{-1}(\varepsilon)\right]} \left(\mu_{\tilde{m}}(x_1) \wedge \mu_{\tilde{\sigma}}(x_2)\right)$$

siendo entonces sus α-cortes, $I_\alpha^\delta$:

$$I_\alpha^\delta = \left\{I = \left[x_1 - x_2\Phi^{-1}(\varepsilon), x_1 + x_2\Phi^{-1}(\varepsilon)\right] \subset R \mid x_1 \in m_\alpha \text{ y } x_2 \in \sigma_\alpha\right\}$$

En cualquier caso, la relación entre ε y δ es, como ya expusimos:

$$\varepsilon = 1 - \frac{1-\delta}{2}$$