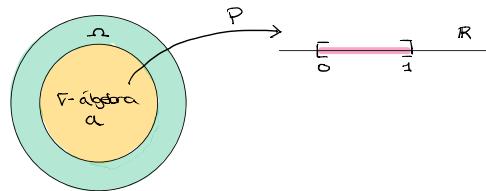
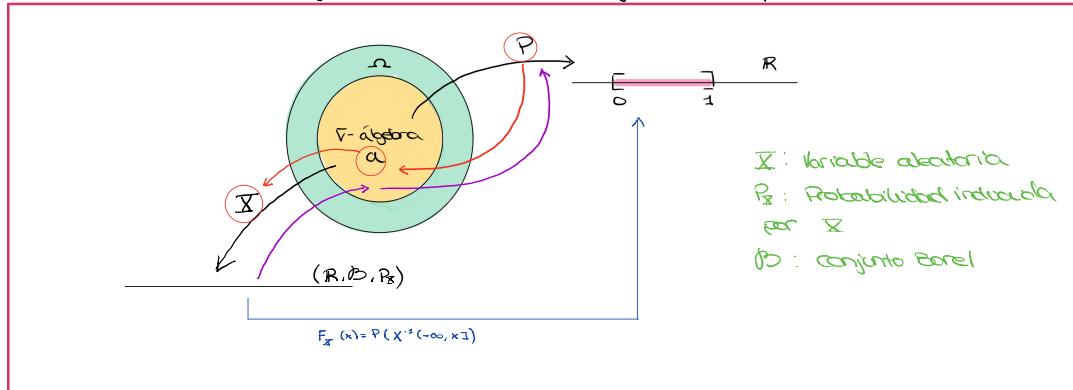


TEMA 5: VARIABLES ALEATORIAS

Hasta ahora nos encontramos en la siguiente situación:



A partir de ahora trabajaremos sobre el siguiente esquema:



X es una **variable aleatoria** si: $\exists: \Omega \rightarrow \mathbb{R}$, función, tal que la imagen inversa de cualquier conjunto Borel es medible, es decir, pertenece al σ -álgebra A .

$$X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}, \forall B \in \mathcal{B} \subseteq \mathcal{G}$$

Las variables aleatorias inducen una medida de probabilidad sobre el espacio Borel

$$P_X: \mathcal{B} \rightarrow \mathbb{R}$$

$$B \rightarrow P_X(B) = P(\omega \in \Omega / X(\omega) \in B)$$

$$(-\infty, x] \rightarrow P_X((-\infty, x]) = P(\omega \in \Omega / X(\omega) \leq x) = F_X(x) \rightarrow \text{Función de distribución}$$

se define la **función de distribución** de una variable aleatoria como:

$$F_X: \mathbb{R} \rightarrow [0, 1]$$

$$F_X(x) = P_X((-\infty, x]) = P(X^{-1}(-\infty, x]) \quad \forall x \in \mathbb{R}$$

Propiedades de la función de distribución:

- Es no decreciente. $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$
- Es continua por la derecha
- $F_X(-\infty) = 0, F_X(+\infty) = 1$

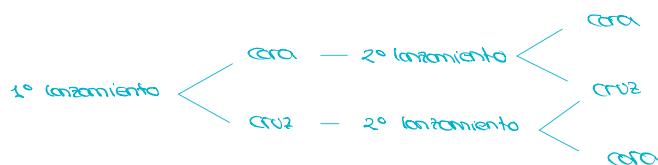
Otras propiedades son:

- El conjunto de puntos de discontinuidad es numerable
- $F_X(x^-) = P[X < x]$
- Solo puede tener discontinuidades de salto y la longitud de salto en un punto de discontinuidad es la probabilidad que toma dicho valor.

$$P[X=x] = P[X \leq x] - P[X < x] = F(x) - F(x^-)$$

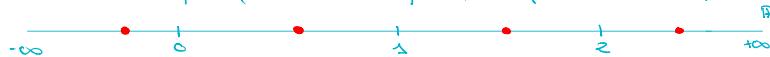
Vamos un ejemplo de variable aleatoria:

Supongamos una moneda y llamemos a X : nº de caras en dos lanzamientos. Sabemos que X es variable aleatoria (a partir de ahora escribiremos v.a.) si $X^{-1}((-\infty, x]) \in \mathcal{A}$ $\forall x \in \mathbb{R}$. Comenzamos distinguiendo casos:



Luego, los posibles casos son que, salga ninguna cara, 1 cara o 2 caras:

$$X^{-1}(2) = \{\text{CC}\}, X^{-1}(1) = \{\text{CT}, \text{TC}\}, X^{-1}(0) = \emptyset$$



$$X^{-1}((-\infty, x]) = \begin{cases} \cdot x < 0 \quad \{\omega \in \Omega / X(\omega) < 0\} = \emptyset \in \mathcal{A} \\ \cdot 0 \leq x \leq 1 \quad \{\omega \in \Omega / X(\omega) \leq x\} = \{\text{CT}, \text{TC}\} \in \mathcal{A} \\ \cdot 1 \leq x < 2 \quad \{\omega \in \Omega / X(\omega) \leq x\} = \{\text{CC}\} \in \mathcal{A} \\ \cdot x \geq 2 \quad \{\omega \in \Omega / X(\omega) \leq x\} = \{\text{CC}, \text{CT}, \text{TC}, \text{CC}\} = \Omega \end{cases}$$

Supongamos ahora el siguiente caso:

$$\mathcal{A} = \{\{\text{CC}\}, \{\text{CT}, \text{TC}\}, \emptyset, \Omega\} \quad \sigma\text{-álgebra}$$

(Ω, \mathcal{A}) y supongamos X : nº de caras $X(\text{CC})=2, X(\text{CT})=X(\text{TC})=1, X(\text{C})=0$

$$X((-\infty, x]) = \begin{cases} \cdot x < 0 \quad \{\omega \in \Omega / X(\omega) \leq x\} = \emptyset \in \mathcal{A}' \\ \cdot 0 \leq x < 1 \quad \{\omega \in \Omega / X(\omega) \leq x\} = \{\text{CT}, \text{TC}\} \notin \mathcal{A} \rightarrow \text{Por lo que } X \text{ no es variable aleatoria.} \\ \cdot \dots \end{cases}$$

Teorema de correspondencia:

Si P es una medida de probabilidad en (Ω, \mathcal{B}) , existe una única función de distribución, F_P en \mathbb{R} , que verifica $F_P(x) = P((-\infty, x]) \quad \forall x \in \mathbb{R}$.

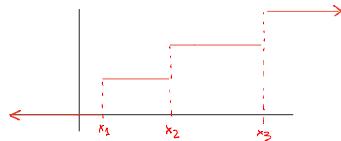
Recíprocamente, si F es una función de distribución en \mathbb{R} , existe una única probabilidad P_F sobre (Ω, \mathcal{B}) que verifica $P_F((-\infty, x]) = F(x) \quad \forall x \in \mathbb{R}$.

Una función de $f: \mathbb{R} \rightarrow \mathbb{R}$, no decreciente, continua a la derecha, con $f(-\infty)=0$ y $f(+\infty)=1$ verifica ser función de una variable aleatoria.

Clasificación de variables aleatorias

Variables aleatorias discretas

Por intuición, X v.a. será discreta si $F(x)$ crece a saltos, es decir, existe, al menos, un conjunto finito numerable de puntos.



Sea X una v.a. discreta con $E = \{x_i / i=1, 2, \dots\}$.

A la función $P: E \rightarrow [0, 1]$ se le denomina función masa de probabilidad:

$$x_i \rightarrow P[X = x_i] = p_i \quad \forall i \in E$$

Esta función verifica $0 \leq P[X = x_i] \leq 1$, ($\forall i \in E$) y $\sum_{i=1}^{\infty} P[X = x_i] = 1$.

La función de distribución viene dado por $F(x) = P[X \leq x] = \sum_{x_i < x} p_i$

Variables aleatorias continua

Por intuición, X v.a. será continua si su función de distribución es absolutamente continua, es decir, existe $f: \mathbb{R} \rightarrow \mathbb{R}$ tal que

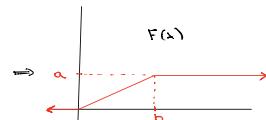
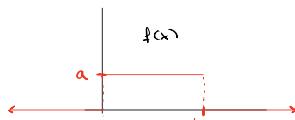
$$F(x) = \int_{-\infty}^x f(t) dt \quad \forall x \in \mathbb{R}$$

esta función recibe el nombre de función de densidad y verifica:

$$\cdot f(x) \geq 0 \quad \forall x \in \mathbb{R}$$

$\cdot f$ es integrable Riemann, ya que tiene al lo sumo un número finito de discontinuidades

$$\cdot \int_{-\infty}^{+\infty} f(x) dx = 1$$



$$F(x) = \int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^b f(x) dx + \int_b^{+\infty} f(x) dx$$

$$= 0 + \int_a^b a dx + 0 = ax \Big|_a^b = ab. \text{ Como:}$$

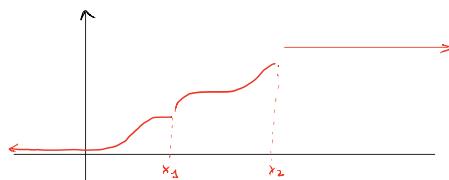
$$F(x) = \int_{-\infty}^{+\infty} f(x) dx = 1 \Rightarrow ab = 1 \Leftrightarrow a = \frac{1}{b}$$

$$\cdot x < 0 \Rightarrow F(x) = \int_{-\infty}^x f(x) dx = 0$$

$$\cdot 0 \leq x \leq b \Rightarrow F(x) = \int_{-\infty}^x f(x) dx = \int_0^x f(x) dx + \int_{-\infty}^0 f(x) dx = \left[\frac{x}{b} \right]_0^x = \frac{x}{b}$$

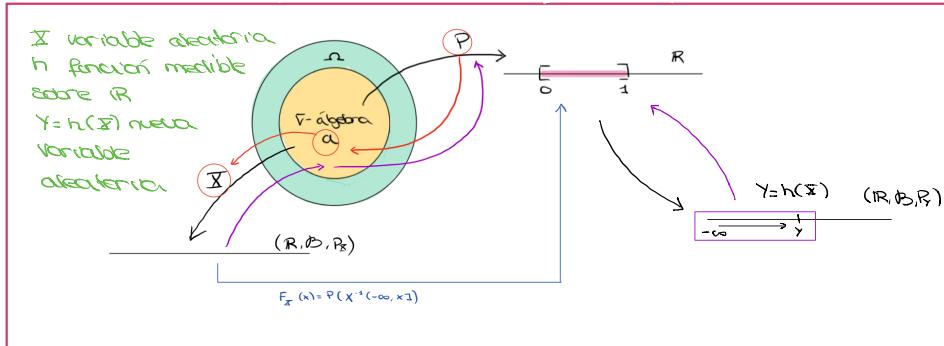
$$\cdot x \geq b \Rightarrow F(x) = \int_{-\infty}^x f(x) dx = \int_b^x f(x) dx + \int_{-\infty}^b f(x) dx = 0 + \left[\frac{x}{b} \right]_b^x = 1$$

Variables aleatorias mixta



Cambio de variable

Actualizamos nuestro espacio original



Teoría general de cambio de variable

Podemos expresar \$Y\$ en función de \$X\$ si \$P_Y(B) = P_X(h^{-1}(B)) \quad \forall B \in \mathcal{F}_Y\$
 $\Leftrightarrow \{ \omega \in \Omega / Y(\omega) \in h^{-1}(B) \}$

Funció n de una variable aleatoria discreta

Sea \$X\$ r.a. discreta

$$E = \{x_1, x_2, \dots\} \subset \mathbb{R}$$

\$h: \mathbb{R} \rightarrow \mathbb{R}\$ medida

\$Y = h(X)\$ r.a. discreta

$$\Rightarrow P[Y=y] = \begin{cases} \sum_{h(x)=y} P[X=x] & y \in h(E) \\ 0 & y \notin h(E) \end{cases}$$

Vamos con un pequeño ejemplo:

x_i	P_i
-1	1/3
0	1/3
1	1/3

E

y_i	P_i
1	1/3
3	1/3
5	1/3

$h(E)$

y_i	P_i
0	1/3
1	2/3

$h(E)$

Funció n de una variable aleatoria continua

En este caso, una función puede dar lugar a una variable aleatoria continua, discreta, continua o mixta.

Continua \$\rightarrow\$ Discreta

\$X\$ r.a. continua

función de densidad \$f

$$A \subset \mathbb{R}$$

\$h: \mathbb{R} \rightarrow \mathbb{R}\$ medida

\$Y = h(X)\$ r.a. discreta

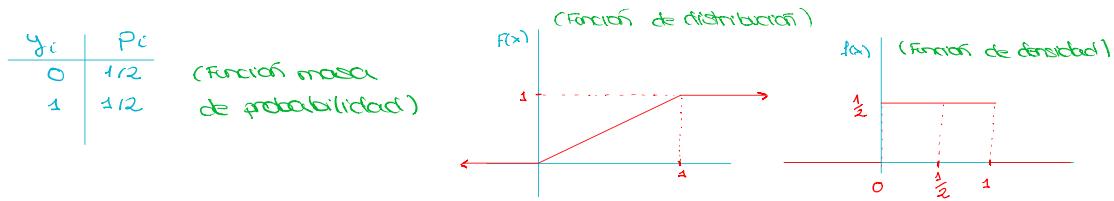
$$\Rightarrow P[Y=y] = \begin{cases} \int_{x \in h^{-1}(y)} f(x) dx & y \in h(A) \\ 0 & y \notin h(A) \end{cases}$$

Supongamos $Y = \begin{cases} 0 & \text{si } x < \frac{1}{2} \\ 1 & \text{si } x \geq \frac{1}{2} \end{cases}$. Calcular la probabilidad $P(Y=0), P(Y=1)$

$$P(Y=0) = \sum_{h(x)=0} P(X < x) = P(X < \frac{1}{2}) = \int_{-\infty}^{1/2} f(x) dx = \int_{-\infty}^0 0 dx + \int_0^{1/2} 1 dx = \frac{1}{2}$$

$$P(Y=1) = \sum_{h(x)=1} P(X > x) = P(X \geq \frac{1}{2}) = \int_{1/2}^{+\infty} f(x) dx = \int_{1/2}^1 1 dx + \int_1^{+\infty} 0 dx = \frac{1}{2}$$

Luego tenemos:



Continua \rightarrow Continua

- X v.a. continua
- f , función de densidad
- $f > 0 \quad \forall x \in [a,b]$
- $h: \mathbb{R} \rightarrow \mathbb{R}$ estrictamente monótona
- Y derivable en $[a,b]$.
- $Y = h(X)$ r.a. continua

$$\Rightarrow g(y) = \begin{cases} f(h^{-1}(y)) |h'(h^{-1}(y))| & y \in h([a,b]) \\ 0 & y \notin h([a,b]) \end{cases}$$

Generalización

$$h(y) = \sum_{k=1}^{\infty} f(h_k^{-1}(y)) \left| \frac{dh_k^{-1}(y)}{dy} \right|$$

Espereanza matemática

Sea X una variable aleatoria discreta con valores $E = \{x_1, x_2, \dots\} \subset \mathbb{R}$, se define la esperanza matemática, media o valor esperado de X , y se denota $E[X]$, como

$$E[X] = \sum_i x_i P[X = x_i]$$

siempre que dicha serie sea absolutamente convergente.

Si la variable aleatoria X es continua con función de densidad f se define la esperanza matemática, media o valor esperado como:

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

siempre que dicha integral sea absolutamente convergente.

En caso de no convergencia absoluta de la serie o de la integral se dice que no existe la esperanza de dicha variable aleatoria.

Propiedades

- La esperanza de una constante $E[c] = c$
- Si una variable aleatoria está acotada, esto es, si $\exists M$ tal que $P[|X| \leq M] = 1 \Rightarrow E[X] \leq M$
- Si $X \geq 0$ y existe su esperanza matemática $\Rightarrow E[X] \geq 0$
- Si X es una variable aleatoria con distribución simétrica respecto del valor c , entonces $\exists E[X]$, su valor será igual a c
- Sea $Y = h(X)$ v.a. discreta, $\Rightarrow \exists E[Y] = \sum_i h(x_i) P[X = x_i] \Leftrightarrow \sum_i |h(x_i)| P[X = x_i] < \infty$
- Sea $Y = h(X)$ r.a. continua $\Rightarrow \exists E[Y] = \int_{-\infty}^{+\infty} h(x) f(x) dx \Leftrightarrow \int_{-\infty}^{\infty} |h(x)| f(x) dx < \infty$

- Linealidad : Si $\exists E[X_i]$, $i=1, 2, \dots, n$, entonces:

$$\exists E \left[\sum_{i=1}^n a_i X_i \right] = \sum_{i=1}^n a_i E[X_i]$$

- Dada X una variable aleatoria y sean $g(x)$ y $h(x)$ dos funciones de X , variables aleatorias cuyas esperanzas existen, entonces:

$$\exists E[g(x)+bh(x)] = aE[g(x)] + bE[h(x)] \quad \forall a, b \in \mathbb{R}$$

- Dada X una aleatoria y sean $g(x)$ y $h(x)$ dos funciones de X , variables aleatorias cuyas esperanzas existen, si $g(x) \leq h(x)$, entonces:

$$E[g(x)] \leq E[h(x)]$$

- La esperanza matemática minimiza el error cuadrático medio:

$$\min_d E[(x-d)^2] = E[(x-E[x])^2]$$

Momentos de una variable aleatoria

Se define el momento no centrado, o centrado en el origen, de orden K de la variable aleatoria X , y lo denominamos como m_K como:

$$m_K = E[X^K], \quad K=1, 2, \dots$$

Se define el momento centrado de orden K de una variable aleatoria X como:

$$\mu_K = E[(X - E[X])^K], \quad K=1, 2, \dots$$

$$\mu_K = \sum_{i=0}^K \binom{K}{i} (-1)^i m_{K-i} m_1^i \quad m_K = \sum_{i=0}^K \binom{K}{i} \mu_{K-i} m_1^i$$

Variancia. Momentos de una variable aleatoria

$$\sigma_X^2 = \text{Var}[X] = \mu_2(X) = E(X - E[X])^2$$

Propiedades

- $\text{Var}[X] = E[X^2] - (E[X])^2$
- Sea X una variable aleatoria cuya varianza existe, entonces $\exists \text{Var}[aX+b] = a^2 \text{Var}[X] \quad \forall a, b \in \mathbb{R}$
- Sea $X \in \mathcal{I}$ variables aleatorias independientes cuyas varianzas existen, entonces $\exists \text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$
- Si x_1, \dots, x_n son variables aleatorias independientes cuyas varianzas existen, entonces $\exists \text{Var}\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 \text{Var}[X_i]$
- La variancia de una variable aleatoria X es cero \Leftrightarrow la variable aleatoria es degenerada o constante.

Medidas de una variable aleatoria

- **Média aritmética**: Discreta: $\sum x_i p_i$ Continua: $\int_{-\infty}^{+\infty} x f(x) dx$
- **Mediana**: $P[X \leq x] = 0,5 \Rightarrow$ Discreta: $\sum p_i = 0,5$, continua: $\int_{-\infty}^x f(x) dx = 0,5$
- **Moda**: x_i con mayor probabilidad. Discreta: $\max\{p_i\}$, continua: $\max\{f(x)\}$
- **Percentil 100**: $P[X \leq x] = n \Rightarrow$ Discreta: $\sum p_i = n$, continua: $\int_{-\infty}^x f(x) dx = n$
- **Recorrido/rango**: $x_K - x_1$
- **Recorrido intercuartílico**: $Q_3 - Q_1$
- **Desviación absoluta respecto la media**: Discreta: $\sum_{i=1}^K |x_i - \bar{x}| p_i$, continua: $|x - \bar{x}| \int_{-\infty}^{\bar{x}} f(x) dx$
- **Desviación absoluta respecto mediana**: Discreta: $\sum_{i=1}^K |x_i - M_0| p_i$, continua: $|x - M_0| \int_{-\infty}^{\bar{x}} f(x) dx$
- **Variancia**: $\sigma_x^2 = m_2 - m_1^2$, Discreta: $m_2 = \sum x_i p_i^2$, $m_1 = \sum x_i p_i$
Continua: $m_2 = \int_{-\infty}^{+\infty} x^2 f(x)^2 dx$, $m_1 = \int_{-\infty}^{+\infty} x f(x) dx$
- **Desviación típica**: $\sqrt{\sigma_x^2}$
- **Coeficiente de asimetría**: $\frac{\max(x_i) - \min(x_i)}{E[x]}$
- **Recorrido relativo**: $\frac{\max(x_i) - \min(x_i)}{Q_3 - Q_1}$
- **Recorrido semi-intercuartílico**: $\frac{Q_3 - Q_1}{Q_3 + Q_1}$
- **Coeficiente de variabilidad de Pearson**: $\frac{\sigma_x}{E[x]}$
- **Índice de dispersión respecto M_e** : $\frac{\sigma_x}{M_e}$
- **Momento central**: $m_k = E[x^k]$
- **Momento no central**: $m_k = E[(x - E[x])^k]$
- **Coeficiente de asimetría de Fisher**: $\frac{\mu_3}{\sigma_x^3}$
- **Coeficiente de asimetría de Pearson**: $\frac{E[x] - M_0}{\sigma_x}$, $\frac{3(M_0 - M_e)}{\sigma_x}$
- **Coeficiente de curtosis de Fisher**: $\frac{\mu_4}{\sigma_x^4} - 3$
- **Coeficiente de curtosis de Kelly**: $K = \frac{1}{2} \cdot \frac{Q_3 - Q_1}{Q_3 + Q_1} - 0,253$

Para percentiles en % u.a.
discreta. Supongamos $P_{60} = M_0$:



Si no cae en un intervalo
cogemos el superior $M_0 = x_2$
Si cae en un intervalo o
también en punto medio
o todo el intervalo



Función generatriz de momentos

Dado una variable aleatoria X , si $\exists t_0 \in \mathbb{R}$ tal que $\forall t \in (-t_0, t_0)$ existe la $E[e^{tX}]$ se dice que existe la función generatriz de momentos y se define como $M_X(t) = E[e^{tX}] \quad \forall t \in (-t_0, t_0)$ (X tiene porque existir, pero si la variable es aleatoria se verifica que existe)

$$\text{Para variable discreta : } M_X(t) = \sum_{i=1}^{\infty} e^{tx_i} P[X = x_i] \quad (\text{Existe si la serie converge})$$

$$\text{Para variable continua : } M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx \quad (M_X(t) < +\infty)$$

Teatrero de unicidad

Si existe la función generatriz de momentos de una variable aleatoria, se verifica que es única y determina de forma única la distribución de la variable. (Una variable aleatoria no puede tener dos funciones generatrizes de momentos ni dos variables con distinta distribución a la misma función generatriz de momentos)

Relación con los momentos

Si $\exists M_X(t)$:

1) \exists todos los momentos

$$2) \forall t \in (-t_0, t_0), M_X(t) = \sum_{n=0}^{\infty} \frac{t^n E[X^n]}{n!}$$

3) Existe todas las derivadas de $M_X(t)$ en un entorno del cero y se verifica:

$$M_X^{(k)}(t) \Big|_{t=0} = E[X^k] \quad \forall k = 1, 2, \dots$$

Otras propiedades

$$\bullet M_X(0) = 1$$

• Sea X una variable aleatoria, con f.o.m. $M_X(t) \quad \forall t \in (-t_0, t_0)$, y sea $Y = ax + b$, $a, b \in \mathbb{R}$, entonces $\exists M_Y(t) = e^{bt} M_X(at)$

Teatrero de Markov

$$\text{Sea } X \text{ v.a. no negativa y } \exists E[X] \Rightarrow P[X \geq \varepsilon] \leq \frac{E[X]}{\varepsilon} \quad \forall \varepsilon > 0$$

$$\text{Discreta: } E[X] = \sum_{x \in E / x < \varepsilon} x P[X=x] + \sum_{x \in E / x \geq \varepsilon} x P[X=x] \geq \sum_{x \in E / x \geq \varepsilon} x P[X=x]$$

$$\sum_{x \in E} x P[X=x]$$

$$E[X] \geq \sum_{x \in E / x \geq \varepsilon} x P[X=x] \geq P[X \geq \varepsilon] = \varepsilon \Rightarrow \frac{E[X]}{\varepsilon} \geq P[X \geq \varepsilon]$$

$$\text{Continua: } E[X] = \int_0^{+\infty} x f(x) dx = \int_0^{\varepsilon} x f(x) dx + \int_{\varepsilon}^{+\infty} x f(x) dx \geq \int_{\varepsilon}^{+\infty} x f(x) dx \geq \varepsilon \int_{\varepsilon}^{+\infty} f(x) dx$$

$$\varepsilon P[X \geq \varepsilon]$$

$$\Rightarrow E[X] \geq \varepsilon P[X \geq \varepsilon] \Rightarrow \frac{E[X]}{\varepsilon} \geq P[X \geq \varepsilon]$$

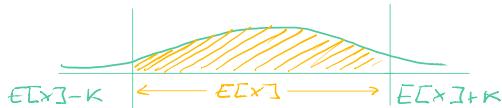
Desigualdad de Chebichev

Si X v.a. / $E[X^2]$ entonces $P(|X - E[X]| \geq K) \leq \frac{Var(X)}{K^2}$

$$P(E[X] - K \leq X \leq E[X] + K) \geq 1 - \frac{Var(X)}{K^2}$$

$$(X - E[X])^2 \geq 0 \quad , \quad \sigma^2 = E[X^2] - (E[X])^2$$

de la
desigualdad
básica
de Markov



$$P(E[X] - K \leq X \leq E[X] + K) \geq 1 - \frac{\sigma^2}{K^2} = 1 - \frac{1}{K^2}$$

Expresiones alternativas:

$$\cdot P(|X - E[X]| < K) \geq 1 - \frac{\sigma^2}{K^2} \quad \forall K > 0$$

$$\cdot P(|X - E[X]| \geq K \sigma_x) \leq \frac{1}{K^2} \quad \forall K > 0$$

$$\cdot P(|X - E[X]| < K \sigma_x) \geq 1 - \frac{1}{K^2} \quad \forall K > 0$$