TEMA 61: Desigualdad de Chebyshev. Coeficiente de variación. Variable normalizada. Aplicación al análisis de los interpretación, comparación de los datos estadísticos

Índice

1. INTRODUCCIÓN

P. Chebyshev fue un célebre matemático ruso del S. XIX, creador de varias escuelas matemáticas en Rusia.
Aunque sus trabajos abordaron diversos campos, su contribución en el campo de la teoría probabilidades destaca sobre las demás.
Así, la conocida como desigualdad de Chebyshev fue fundamental para demostrar la ley débil de los grandes números, así como, para estimar la cantidad de primos menores o iguales a $n \in \mathbb{Z}^+$.

Por otro lado, sus discípulos más famosos: Markov y Liapunov, siguieron sus trabajos ampliándolos y mejorándolos.
En este tema, además, abordaremos los conceptos de: coeficiente de variación y variable normalizada, dando algunas de sus aplicaciones prácticas en la actualidad.

2. DESIGUALDAD DE CHEBYSHEV

En esta sección, consideraremos que todas las v.a. son continuas. Aunque, la demostración para el caso discreto es muy parecido.
Además, la hipótesis de función medible, garantizara que el conjunto de la transformación se puede "medir", es decir, calcular su probabilidad.

TEOREMA (Desigualdad de Markov) Sea X una variable aleatoria (v.a) ~~entonces~~ y $g(x)$ una función medible no negativa, entonces

$$\forall r>0 \qquad P(g(x)>r) \leq \frac{E[g(x)]}{r}$$

D Supongamos $g(x) \geq 0 \quad \forall x \in \mathbb{R}$, $r>0$

Entonces,
$$S := \{x \in \mathbb{R} : g(x) > r\}$$

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x) f(x)\,dx = \int_S g(x) f(x)\,dx + \underbrace{\int_{S^c} g(x) f(x)\,dx}_{\text{no negativo}}$$

$$\underset{\substack{g(x)\geq 0 \\ \forall x \in \mathbb{R}}}{\geq} \int_S g(x) f(x)\,dx \underset{\substack{g(x)>r \\ \forall x \in S}}{\geq} r \int_S f(x)\,dx = r\,P(g(x)>r)$$

$$\Rightarrow \quad P(g(x)>r) \leq \frac{E[g(x)]}{r} \qquad \square$$

Veamos ahora como este resultado, demostrado por Markov, permite deducir otros muy interesantes

Corolario 1 - Bajo las condiciones del teorema anterior se puede deducir que $\forall \varepsilon > 0$

$$P(|X - E[X]| > \varepsilon) \leq \frac{Var[X]}{\varepsilon^2}$$

$\boxed{D}$ Sea $g(x) := |x - E[x]| \geq 0$ $\forall x \in \mathbb{R}$ y medible

Entonces, dado que

$$\forall \varepsilon > 0 \quad |x - E[x]| > \varepsilon \iff \underbrace{|x - E[x]|^2}_{(x-E[x])^2} \geq \varepsilon^2$$

$$P(|X - E[X]| > \varepsilon) = P((X - E[X])^2 > \varepsilon^2)$$

Por el teorema anterior, $\cancel{\varepsilon > 0}$

$$P(|X - E[X]|^2 > \varepsilon^2) \leq \frac{E[(X - E[X])^2]}{\varepsilon^2}$$

$$\Rightarrow \quad P(|X - E[X]| > \varepsilon) \leq \frac{Var[X]}{\varepsilon^2}. \quad \square$$

Corolario 2 - $\forall k \in \mathbb{N}$. Suponiendo las hipótesis del T^a anterior se deduce que $\forall \varepsilon > 0$

$$P(|X \cancel{- E[X]}| > \varepsilon) \leq \frac{E[|X|^k]}{\varepsilon^k}$$

$\boxed{D}$ Sea $g(x) := |x|^k \geq 0$, medible

Entonces $|x| > \varepsilon \iff |x|^k > \varepsilon^k$

$P(|X|^k > \varepsilon^k) = P(|X| > \varepsilon)$. Si aplicamos el T^a para $g(x)$

$$P(|X| > \varepsilon) = P(|X|^k > \varepsilon^k) \leq \frac{E[|X|^k]}{\varepsilon^k} \quad \square$$

Corolario 3: (Desigualdad de Chebyshev)

Sea X una v. a. entonces: Si $K > 0$

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

ii) $P(|X - E[X]| < K\sigma) \geq 1 - \frac{1}{K^2}$

$\boxed{D}$ Sea $g(x) = |X - E[X]| \geq 0 \quad \forall x \in \mathbb{R}$, medible

i) Si $\sigma = +\sqrt{Var(X)}$ entonces si $K > 0$

$\Rightarrow K\sigma > 0$. Aplicamos la D. Markov

~~$P(|X-E[X]| \geq K\sigma) \leq$~~

Corolario 1

$$P(|X - E[X]| \geq K\sigma) = P(|X - E[X]|^2 \geq K^2\sigma^2) \leq$$

$$\leq \frac{Var[X]}{K^2\sigma^2} = \frac{\cancel{\sigma^2}}{K^2\cancel{\sigma^2}} = \frac{1}{K^2}$$

ii) Si sabemos que ii) es el complementario de i)

$$P(|X - E[X]| \geq K\sigma) = 1 - P(|X - E[X]| < K\sigma)$$

solo hay que aplicar propiedades de las desigualdades.

Veamos algunas aplicaciones interesantes de las propiedades anteriores

Sea X una v. a. entonces por el Corolario 3

Sabemos que

$$P(|X-E[X]| < k\sigma) > 1-\frac{1}{k^2}$$

Luego si $\boxed{k=2}$, entonces

$$|X-E[X]| < 2\sigma \iff$$

$$\iff -2\sigma < X-E[X] < 2\sigma \iff$$

$$\iff E[X]-2\sigma < X < E[X]+2\sigma$$

Si aplicamos que

$$P(|X-E[X]| < 2\sigma) > 1-\frac{1}{4} = \frac{3}{4}$$

$$\Rightarrow P(E[X]-2\sigma < X < E[X]+2\sigma) > \frac{3}{4}$$

Si aplicamos para $\boxed{k=3}$ entonces

$$P(E[X]-3\sigma < X < E[X]+3\sigma) > 1-\frac{1}{9} = \frac{8}{9}$$

<u>Ejemplo</u>.- Sea $X := \{$ ~~clientes~~ clientes que llega a una tienda $\}$

Si $E[X]=20$, $\sigma = 2$

$$P(16 < X < 24) = P(20-4 < X < 20+4)$$

$$= P(20-2\sigma < X < 20+2\sigma) > \frac{3}{4}$$

3. COEFICIENTE DE VARIACIÓN

Sea X una v.a tal que

$E[X] = \mu$; donde la desviación típica σ, indica la dispersión entre los valores de X con respecto a la media μ

$Var[X] = \sigma^2$

Habitualmente nos encontramos el problema de comparar dos magnitudes X e Y de forma correcta. Así, puesto que la desv. típica, la media y la varianza dependen de las ~~media~~, unidades de medida, ~~debemos aplicar un corrector para tratar de unificarlas.~~ introduciremos el coeficiente de variación de Pearson para tratar de soslayar el problema.

Def^n.- Sea X una v.a, con $E[X] = \mu$ y $Var[X] = \sigma^2$, entonces el coeficiente de variación de Pearso viene dado por

$$\boxed{CV = \frac{\sigma}{\mu}}$$

Observación.- Sea X una v.a con μ, σ^2.

Entonces, $Y = aX + b$ con $a, b \in \mathbb{R}$ satisface que $E[Y] = a\mu + b$ y $Var[Y] = a^2\sigma^2$

Luego, $CV_X = \frac{\sigma}{\mu}$ $\qquad \frac{a\sigma}{a\mu + b} = CV_Y$

Luego, $CV_X = CV_Y \iff b = 0$

Así, si hay un cambio de escala únicamente el CV es el mismo. Sin embargo, esta medida es sensible a un cambio de localización, por tanto, en el análisis de los datos hay que ser cuidadosos con dicha característica

4. VARIABLE NORMALIZADA

En la sección anterior hemos introducido una medida para comparar datos a diferentes escalas. A continuación, presentamos otra técnica muy habitual y útil ...

Def.^{n}- Una v.a. X se dice tipificada o normalizada o estandarizada sii
$$E[X] = 0$$
$$Var[X] = 1$$

Proposición.- Sea X una v.a con $E[X] = \mu$ y $Var[X] = \sigma^2$. Entonces

$$Y = \frac{X-\mu}{\sigma} \text{ está enstandarizada}$$

$\boxed{D}$
$$E[Y] = \frac{1}{\sigma} E[X-\mu] = \frac{1}{\sigma}(E[X]-\mu) =$$
$$= \frac{1}{\sigma}(\mu-\mu) = 0$$
$$Var[Y] = \frac{1}{\sigma^2} Var[X-\mu] = \frac{Var[X]}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1$$

$\square$

5. APLICACIONES AL ANALISIS DE LOS DATOS

Como hemos mencionado en las secciones anteriores el CV nos permite hacer un estudio obviando las diferentes escalas, siempre que no haya también un cambio de localización. Así, esta medida es invariante respecto a las escalas.

En cambio, ¿qué ocurre con las estandarización?

Sea X una v.a. tal que $\begin{cases} E[X] = \mu \\ Var[X] = \sigma^2 \end{cases}$

$$Z = \frac{X - \mu}{\sigma^2} \text{ está estandarizada}$$

Consideremos $Y = aX + b$, $a, b \in \mathbb{R}$ una transformación lineal de X, ¿qué ocurre si la estandarizamos?

$$T = \frac{Y - E[Y]}{\sqrt{Var[Y]}} = \frac{Y - (aE[X] + b)}{\sqrt{a^2 Var[X]}} =$$

$$= \frac{aX + b - a\mu - b}{\sqrt{a^2 \sigma^2}} = \frac{\cancel{a}(X - \mu)}{\cancel{a}\sigma} = Z$$

Luego, si estandarizamos $\left.\begin{matrix} X \\ Y \end{matrix}\right\} \to Z \Rightarrow$

Esto es, ~~toda transformación lineal~~ la estandarización es INVARIANTE sobre transformaciones lineales. Así, estandarizar un v.a es muy frecuente con el objetivo de eliminar problemas de escalas y localizaciones

Nota pedagógica: Todos los procesos de análisis de datos complejos, como los que se usan para generar modelos de IA, estandarizan los datos para evitar problemas numéricos como:

1. parámetros numéricos grandes
2. distorsiones de impacto e importancia
3. etc

6. CONCLUSIONES

Aunque es un tema complejo, este tema se puede introducir en el aula cuando en Bachillerato se trabaja la distribución Normal. Así una actividad interesante podría ser observar si se cumple la desigualdad de Chebyshev cuando $X \sim N(0,1)$ y por tanto ¿cuál es

$$P(-2 < X < 2)?$$

Además, humanizar las clases introduciendo personajes históricos como Chebyshev es la mejor forma de acercar las matemáticas a nuestros alumnos.

1h 15'