ESTAD_T4. DISTRIB. UNIDIMENSIONALES. ESPERANZA MATEMÁTICA. PROPIEDADES. MOMENTOS de una V.A. UNIDIM. OTRAS MEDIDAS de DISP, POSIC y FORMA. TEOREMA de MARKOV y desig. TCHEBYCHEV.

1- DISTRIB. UNIDIMENSIONALES.

Recordemos la definición de variable aleatoria.

$\xi : \Omega \longrightarrow \mathbb{R}$

$s \longmapsto (-\infty, x]$ es una v.a. $\Leftrightarrow \xi^{-1}(-\infty, x] \in \Omega$.

Una variable aleatoria es una aplicación que va desde el espacio de probabilidad a la recta real (con σ-álgebra de Borel), de tal modo que la antiimagen de cada boreliano (intervalo) es un suceso.
La variable aleatoria es una transformación del espacio de sucesos posibles, no medible, a la recta real, sí medible.

Una variable aleatoria, representativa de un experimento aleatorio, queda definida cuando conocemos su campo de variación y el conjunto de probabilidades con que toma valores en ese campo, es decir, su distribución de probabilidad.
Pero muchas veces no es posible (suma de uniformes) ó no hace falta conocer la distrib. de probabilidad de una v.a. En esas ocasiones, resulta interesante conocer una serie de características que nos den información sobre determinados ~~la v.a.~~ aspectos de la v.a., como pueden ser su posición, dispersión o forma. Las características más conocidas son la esperanza, μ y la varianza, σ^2.

ESTAD_T4.

---

2 – ESPERANZA MATEMÁTICA. PROPIEDADES.

Aunque el término de esperanza comenzó a utilizarse relacionado a los juegos de azar (concepción clásica de la probabilidad), la interpretación de la esperanza matemática como el valor medio de la distribución de probabilidad, es el plenamente aceptado con toda generalidad.

Entender la esperanza matemática, $E[\xi]$ como el valor al que tendería la media aritmética $\bar{x}$ con un nº sufic. grande de observaciones pasa por admitir la definición de probabilidad como límite de la frecuencia relativa asociada a cada alternativa (concepción frecuencialista)

$$E[\xi] = \sum x_i p_i \text{ , donde } p_i = \lim_{N \to \infty} \frac{n_i}{N} \quad \sim \quad \bar{x} = \sum x_i \frac{n_i}{N}$$

$$\cancel{\bar{x} = \sum x_i \frac{n_i}{N}}$$

La definición formal de esperanza matemática requiere (la formulación matemática del concepto de esperanza de nuevo) la distinción entre v.a. de tipo discreto y de tipo continuo.

$$E[\xi] = \begin{cases} \sum_i x_i p_i \text{ ,} & \xi \text{ v.a. discreta} \\ \int_{-\infty}^{+\infty} x f(x)\,dx \text{ ,} & \xi \text{ v.a. continua} \\ \sum_{i \in D} x_i p_i + \int_{x \in C} x f(x) \text{ .} & \xi \text{ v.a. mixta } \begin{cases} D \to \text{campo discreto} \\ C \to \text{campo continuo} \end{cases} \end{cases}$$

Hablar de esperanza matemática, es lo mismo que hablar de valor medio, media, valor esperado, valor probable ó sencillamente esperanza.

<u>OJO</u>: la esperanza matemática <u>no</u> siempre existe. Dependiendo ~~del tipo~~ de v.a., su existencia depende de la ~~convergencia de una~~ (absoluta) serie infinita ó de una integral impropia.

ξ v.a. discreta, $\exists E[\xi]$ si $\sum_{k=1}^{\infty} |x_k| p_k < +\infty$

ξ v.a. continua, $\exists [\xi]$ si $\int_{-\infty}^{+\infty} |x| \cdot f(x)\, dx < +\infty$. absol. conv.

PROPIEDADES de la ESPERANZA

P1 — La esperanza de una constante es igual a la misma constante. (C $\to$ distrib. degenerada ó causal).

$$E[C] = C.$$

<u>Dem</u>: C v.a. discreta, $E[C] = \sum x_i p_i = C \cdot 1 = C$.

P2 — La esperanza de la suma algebraica de v.a. es igual a la suma algebraica de las esperanzas de cada una de las v.a.

$$E[\xi_1 \pm \xi_2 \pm \ldots \pm \xi_n] = E[\xi_1] \pm E[\xi_2] \pm \ldots \pm E[\xi_n].$$

<u>Dem</u>: Se demuestra para $n=2$ y se generaliza por inducción.

ξ_1, ξ_2 v.a. discretas:

probab. conjunta

$$E[\xi_1 \pm \xi_2] = \sum_i \sum_j (x_i \pm y_j) p_{ij} = \sum_i \sum_j x_i p_{ij} \pm \sum_i \sum_j y_j p_{ij} =$$

probab. marginal

$$= \sum_i x_i \left[\sum_j p_{ij}\right] \pm \sum_j y_j \left[\sum_i p_{ij}\right] = \sum_i x_i p_{i\cdot} \pm \sum_j y_j p_{\cdot j}$$

$$= E[\xi_1] \pm E[\xi_2].$$

ξ_1, ξ_2 v.a. continuas:

f.d. conjunta

$$E[\xi_1 \pm \xi_2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \pm y) f(x,y)\, dx\, dy = \ldots$$

$$\ldots = \iint_{-\infty}^{+\infty} x\, f(x,y)\, dx\, dy \pm \iint_{-\infty}^{+\infty} y\, f(x,y)\, dx\, dy =$$

$$= \int_{-\infty}^{+\infty} x \Big(\underbrace{\int_{-\infty}^{+\infty} f(x,y)\, dy}_{\text{f.d. marginal de } x} \Big) dx \pm \int_{-\infty}^{+\infty} y \Big(\underbrace{\int_{-\infty}^{+\infty} f(x,y)\, dx}_{\text{f.d. marginal de } y} \Big) dy =$$

$$= \int_{-\infty}^{+\infty} x \cdot f_1(x)\, dx \pm \int_{-\infty}^{+\infty} y \cdot f_2(y)\, dy = E[\xi_1] \pm E[\xi_2].$$

P3_ La esperanza del producto de v.a. es igual al producto de las esperanzas de cada una de las v.a. si y sólo si son estad. independientes

$E[\xi_1 \cdot \ldots \cdot \xi_n] = E[\xi_1] \cdot E[\xi_2] \ldots E[\xi_n] \Leftrightarrow \xi_i$ indep.

(En el caso de que no sean estad. independientes, la esperanza se calcula direct. a través de la definición, no se puede separar)

Dem.: Para $n=2$, para $n>2$ por inducción.

ξ_1, ξ_2 v.a. discretas:

$$E[\xi_1 \cdot \xi_2] = \sum_i \sum_j x_i \cdot y_j\, p_{ij} \overset{\xi_1 \text{ indep } \xi_2 \Leftrightarrow p_{ij} = p_{i\cdot} \cdot p_{\cdot j}\ \forall i,j}{=} \sum_i \sum_j x_i \cdot y_j \cdot p_{i\cdot} \cdot p_{\cdot j} =$$

$$= \sum_i x_i\, p_{i\cdot} \cdot \sum_j y_j\, p_{\cdot j} = E[\xi_1] \cdot E[\xi_2].$$

ξ_1, ξ_2 v.a. continuas:

$$E[\xi_1 \cdot \xi_2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x\, y\, f(x,y)\, dx\, dy \overset{\xi_1 \text{ indep } \xi_2 \Leftrightarrow f(x,y) = f_1(x) f_2(y)}{=} \iint_{-\infty}^{+\infty} x\, y\, f_1(x) f_2(y)\, dx\, dy =$$

$$= \int_{-\infty}^{+\infty} x\, f_1(x)\, dx \int_{-\infty}^{+\infty} y\, f_2(y)\, dy = E[\xi_1]\, E[\xi_2].$$

P4_ La esperanza de las desviaciones de los valores de la v.a. respecto a su media es cero.

$$E[\xi - E[\xi]] = E[\xi - \mu] = 0.$$

↳ $E[\xi] \equiv$ centro de gravedad de la distrib.

↳ Parámetro de tendencia central de la distrib.

Dem: $E[\xi - \mu] \overset{P2}{=} E[\xi] - E[\mu] \overset{P1}{=} E[\xi] - \mu = \mu - \mu = 0$.

P5 – Si a una v.a. se le suma una constante, su esperanza queda modificada en esa misma constante.

$$E[\xi + C] = E[\xi] + C$$

Dem: $E[\xi + C] \overset{P2}{=} E[\xi] + E[C] \overset{P1}{=} E[\xi] + C$.

P6 – Si una v.a. se le multiplica una constante, su esperanza matemática también queda multiplicada por esa cte.

$$E[C \cdot \xi] = C \cdot E[\xi]$$

Dem: $E[C \cdot \xi] \overset{P3 \text{ (siempre indep)}}{=} E[C]\, E[\xi] \overset{P1}{=} C \cdot E[\xi]$.

(P5+P6) $\Longrightarrow$ Corolario: La esperanza matemática de una transformación lineal de una v.a. será la transformación lineal de la esperanza matemática de la v.a.

$$E[a + b\xi] = a + b\,E[\xi]$$

Esperanza de una función de ξ, $E[g(\xi)]$.

$E[g(\xi)] = \sum_i g(x_i)\, p_i$ siempre que $\sum_i |g(x_i)|\, p_i < +\infty$. ξ v.a. DISCR.

$E[g(\xi)] = \int_{-\infty}^{+\infty} g(x) \cdot f(x)\, dx$ siempre que $\int_{-\infty}^{+\infty} |g(x)|\, f(x)\, dx < +\infty$ ξ v.a. CONT.

Así, puede determinarse la esperanza de la transformación $g(\xi)$ a través de la distrib. de probab. de la v.a. ξ, sin necesidad de obtener la distrib. de probab. de la transformación.

Las propiedades vistas anteriormente son extensivas a la transformación $g(\xi)$.

3. MOMENTOS de una V.A. UNIDIMENSIONAL

a) momentos respecto al origen

Momento respecto al origen de orden r, α_r, de la variable aleatoria ξ es la esperanza de la v.a. elevada a r.

$$\alpha_r = E[\xi^r] = \begin{cases} \sum_i x_i^r p_i & \text{si } \sum_i |x_i^r| p_i < +\infty \text{ para } \xi \text{ v.a. discreta} \\ \int_{-\infty}^{+\infty} x^r f(x)dx & \text{si } \int_{-\infty}^{+\infty} |x^r| f(x)dx < +\infty \text{ para } \xi \text{ CONT.} \end{cases}$$

Teorema: $\exists \alpha_s \Rightarrow \exists \alpha_t, \forall t < s. \Rightarrow \exists \mu_t \ \forall t < s$.

Normalmente se trabaja con $\alpha_1, \alpha_2, \alpha_3$ y α_4.

$\alpha_0 = 1$, $\alpha_1 = E[\xi] = \mu$.

b) momentos respecto a la media

Momento de orden r respecto a la media, μ_r, es la esperanza de la variable menos su media a la r-ésima potencia.

$$\mu_r = E[(\xi-\mu)^r] = \begin{cases} \sum_i (x_i-\mu)^r p_i & \text{si } \sum |(x_i-\mu)^r| p_i < +\infty \ \ \xi \text{ DISCR.} \\ \int_{-\infty}^{+\infty} (x-\mu)^r f(x) dx & \text{si } \int_{-\infty}^{+\infty} |(x-\mu)^r| f(x)dx < +\infty \ \ \xi \text{ CONT.} \end{cases}$$

$\mu_1 = 0$ (por P4), $\mu_2 = \sigma^2$, μ_3, μ_4.

Teorema: Todo momento ~~de orden r~~ respecto a la media se puede calcular a partir de los momentos respecto al origen.

$$\mu_r = E[\xi-\mu]^r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \mu^k \alpha_{r-k} \quad // \ \mu \equiv \text{media}$$

Para $r=2$, $\mu_2 = \alpha_2 - \alpha_1^2$

$r=3$, $\mu_3 = \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3$

$r=4$, $\mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4$

4_ OTRAS MEDIDAS de POSICIÓN, DISPERSIÓN y FORMA.

a) de POSICIÓN:

Moda: Valor más probable → el que presenta la máxima probabilidad.

$$Mo = \begin{cases} x_i \,/\, P_i \geq P_j \quad \forall i \neq j \text{ en } \xi \text{ discreta} \\ x_0 \,/\, f(x_0) = \text{máx } f(x) \; (f'(x_0)=0,\, f''(x_0)<0). \end{cases}$$

La moda de una distrib. no tiene por qué ser única ni es necesariamente un valor de tendencia central.

Mediana: Valor de la variable situado en la mitad de la distribución. Deja a cada lado 1/2 de la masa de probabilidad.

$$P(\xi \leq Me) \geq \frac{1}{2} \quad y \quad P(\xi \geq Me) \geq 1/2$$

Cuantiles: Un cuantil de orden p es un valor de la distribución que deja a su izquierda la probab. acumulada p.

$$P(\xi \leq x) \geq p \quad y \quad P(\xi \geq x) \geq 1-p \quad \text{para } 0<p<1.$$

Casos particulares de cuantiles son los cuartiles, que dividen la distrib. en 4 partes iguales, los deciles y los percentiles.

b) de DISPERSIÓN:

Dispersión es la mayor o menor variabilidad de los valores de la variable alrededor de su valor medio.

Como $E[\xi - \mu] = 0$, por ser μ el centro de gravedad de la distribución (los signos positivos y los signos negativos de las desviaciones se compensan), para eliminar el signo se toman valores absolutos ó potencias pares

Desviación media respecto a la media

$$D(\xi) = E[|\xi - \mu|]$$

Varianza

$$V(\xi) = \sigma^2 = \mu_2 = E[(\xi-\mu)^2]$$ disc. ← la + utilizada; cont.

P1 _ $V(\xi) \geq 0$.

P2 _ $ECM(\xi) = E[\xi - k]^2$ es mínimo cuando $k = \mu$

$$\min_k ECM(\xi) = \min_k E(\xi - k)^2 = E[\xi-\mu]^2 = V(\xi).$$

P3 _ $V(\xi) = \sigma^2 = \mu_2 = \alpha_2 - \alpha_1^2$

P4 _ $V(\xi_1 \pm \xi_2) = V(\xi_1) + V(\xi_2) \pm 2\,cov(\xi_1, \xi_2)$

P5 _ $V(\xi \pm c) = V(\xi)$

P6 _ $V(c\xi) = c^2 V(\xi)$

$\rightarrow V(a + b\xi) = b^2 V(\xi)$.

Desviación típica

$\sigma = +\sqrt{\sigma^2} = +\sqrt{E[\xi-\mu]^2}$ definida en las mismas unidades de medida que la v.a. y su media.

Tipificación de una v.a.

Sea ξ una v.a. tq $E[\xi] = \mu$, $V[\xi] = \sigma^2$

la v.a. tipificada resulta de restar la media y después dividir por la desv. típica, por lo que será otra v.a. con esperanza 0 y varianza 1.

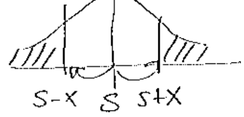
$$\xi^* = \frac{\xi - \mu}{\sigma} \quad tq \quad E[\xi^*] = 0, \; V[\xi^*] = 1.$$

c) de FORMA:

c.1. Asimetría → falta de simetría.

ξ es simétrica respecto a un eje perpendicular a OX en el punto s si

$$P(\xi \geq s + x) = P(\xi \leq s - x)$$



Lo habitual es considerar $s = \mu$.

Para medir la simetría lo lógico es utilizar μ_r.

$\mu_1 = 0$ por P3 de la media.

$\mu_2 \geq 0$ siempre

μ_3 es el 1er momento con signo $\begin{cases} \mu_3 = 0 \rightarrow \text{simetría} \\ \mu_3 > 0 \rightarrow \text{asim. } (+) \\ \mu_3 < 0 \rightarrow \text{asim. } (-) \end{cases}$

(Fisher) elaboró un coeficiente de asimetría adimensional e invariante a cambios de origen y de escala basado en μ_3.

$$\gamma_1 = \frac{\mu_3}{\sigma^3}$$

c.2. Curtosis → apuntamiento respecto a una distrib. Normal de una distrib. campaniforme y aprox. simétrica.

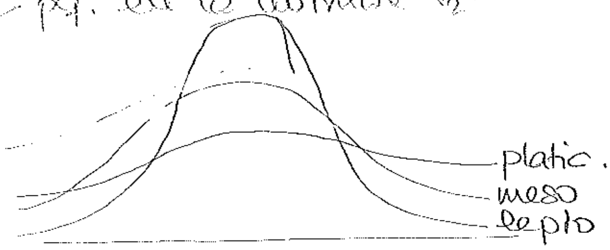
(Fisher) propuso como coeficiente de curtosis

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \quad \leftarrow \text{pq. en la normal } \beta_2 = 3$$

$\gamma_2 = 0 \rightarrow$ mesocúrtica

$\gamma_2 > 0 \rightarrow$ leptocúrtica

$\gamma_2 < 0 \rightarrow$ platicúrtica



5. TEOREMA de MARKOV y desig. TCHEBYCHEV.

Teorema: Dada ξ v.a. y $g(\xi)$ una transformación tq $g(\xi) \geq 0$. Para una cte positiva $K > 0$ se verifica

$$P[g(\xi) \geq K] \leq \frac{E[g(\xi)]}{K}$$

Dem: Dividimos el campo de

Dem: Dividimos el campo de variación de ξ en dos subcjtos,

$$\underbrace{\quad g(\xi) < K \quad}_{} \underset{K}{|} \underbrace{\quad g(\xi) \geq K \quad}_{}$$

Entonces la esperanza de la transformación

$$E[g(\xi)] = \int_{-\infty}^{+\infty} g(x) f(x)\,dx = \int_{g(\xi)\geq K} g(x) f(x)\,dx + \int_{g(\xi)<K} g(x) f(x)\,dx \geq$$

$$\geq \int_{g(\xi)\geq K} g(x) f(x)\,dx \geq \int_{g(\xi)\geq K} K f(x)\,dx =$$

$$= K \int_{g(\xi)\geq K} f(x)\,dx = K \cdot P[g(\xi)\geq K].$$

luego $E[g(\xi)] \geq K \cdot P[g(\xi) \geq K] \Rightarrow$ ($K>0$ no cambia sentido desig.)

$$P[g(\xi)\geq K] \leq \frac{E[g(\xi)]}{K} \quad \text{cqd}$$

Desigualdad de Chebychev

Es un caso particular del teorema de Markov con la transformación $g(\xi) = (\xi-\mu)^2$

Así, $E[g(\xi)] = E[(\xi-\mu)^2] = \sigma^2$

y el tma resulta

$$P[(\xi-\mu)^2 \geq K] \leq \frac{\sigma^2}{K}$$

ó bien

$$P[|\xi-\mu| \geq \sqrt{K}] \leq \frac{\sigma^2}{K} \leftarrow \text{desig. chebychev.}$$

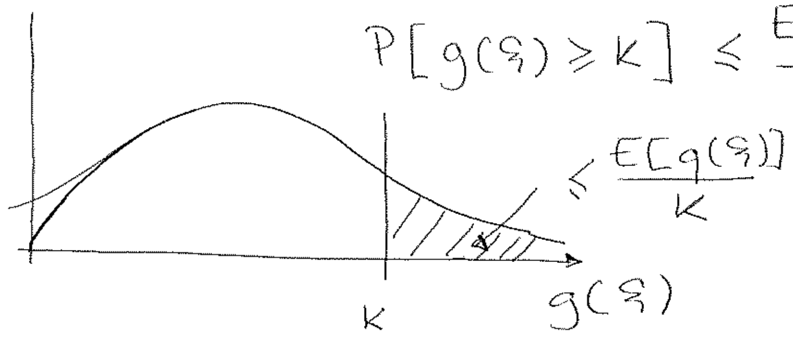
ó bien

$$P[|\xi-\mu| \geq K\sigma] \leq \frac{1}{K^2} \sim P[|\xi-\mu| < K\sigma] \geq 1 - \frac{1}{K^2}.$$

Tma. Markov : ξ v.a. $g(\xi) \,/\, g(\xi) \geq 0$

Dada una cte $K > 0$ se verifica

$$P[g(\xi) \geq K] \leq \frac{E[g(\xi)]}{K}$$



La cota de probabilidad es inversamente proporcional al valor inferior de la transformación ($E[g(\xi)]$ fija)

Al ser $K>0$ se puede decir que $K \uparrow \Rightarrow$ Probab $\downarrow$ SIEMPRE

La gracia es el ritmo de decrecimiento.

Desig. Chebichev :

Es una particularización del Tma de Markov, para la función $g(\xi) = (\xi - \mu)^2$, que es no negativa.

$$E[g(\xi)] = E[(\xi-\mu)^2] = V(\xi) = \sigma^2$$

Por lo que la desig. de Chebichev relaciona la media y la varianza de una v.a., de la cual no tenemos por qué conocer su distrib. de probabilidad.

$$P[(\xi-\mu)^2 \geq K] \leq \frac{V(\xi)}{K} \longrightarrow P[|\xi - \mu| \geq \sqrt{K}] \leq \frac{\sigma^2}{K}$$

que es lo mismo que

$$P[-\sqrt{K} \leq \xi - \mu \leq +\sqrt{K}] \leq \frac{\sigma^2}{K} \longrightarrow P[\mu - \sqrt{K} \leq \xi \leq \mu + \sqrt{K}] \leq \frac{\sigma^2}{K}$$

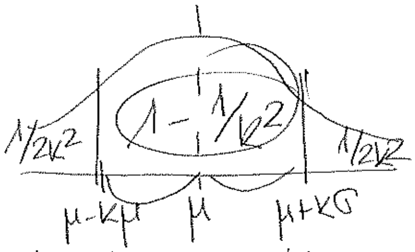
por lo que la desig. de Chebichev proporciona una cota de probabilidad para un I.C. de la v.a. centrado en su media y con una desviación de $\sqrt{K}$. Dicha cota de probab. depende de la varianza de ξ y de la cte K.

Una versión muy utilizada de la desig. de Chebichev es aquella que deja la cota de probab. en función exclusivamente de la cte K.

Desig. → $P[|\xi-\mu| \geq \sqrt{K}] \leq \frac{\sigma^2}{K}$ $\cancel{P[|\xi}$

Si hacemos $\sqrt{K} = K\sigma$
$\uparrow$ cte de

$$P[|\xi-\mu| \geq K\sigma] \leq \frac{\cancel{\sigma^2}}{K^2\cancel{\sigma^2}} = \frac{1}{K^2}$$



Ahora el intervalo de confianza sigue centrado en μ, pero su error depende de la cte y de σ

I.C. → $\mu \pm K\sigma$ $\qquad P(\xi \notin I.C.) \leq \frac{1}{K^2}$

Para el complementario $\qquad P(\xi \in I.C.) \geq 1 - \frac{1}{K^2}$

$$P[|\xi-\mu| < K\sigma] \geq 1 - \frac{1}{K^2}$$

Consecuencias:

1) La desig. confirma que σ^2 es una buena medida de dispersión.

↳ A medida que k crece, la probab. de que ξ tome valores fuera del intervalo $(\mu \pm K\sigma)$ es cada vez más pequeña.

↳ Para una misma cota de probabilidad, la amplitud del intervalo dependerá de σ, es decir, dependerá de dispersión de la v.a.

2) La desig. de Chebichev resulta muy útil en los casos en que la distrib. de ξ es desconocida (⇒ no se pueden calcular las probab. directamente), de tal modo que conocidas únicamente la μ y la σ^2 de una distrib, se pueden determinar cotas de probabilidad para intervalos de la forma $(\mu \pm K\sigma)$