TEMA VI: VARIÁBEIS ALEATÓRIAS

Estudo de fenómeno aleatório	EX POST *Resultados cuantificábeis* (despois da sua execución)	VARIÁBEL ESTATÍSTICA
	EX ANTE *Asócia-se sucesos posíbeis a números reais mediante* (antes da sua execución)	VARIÁBEL ALEATÓRIA

Consideramos un experimento aleatório con espazo de probabilidade asociado (Ω, A, P):

• **VARIÁBEL ALEATÓRIA** é unha aplicación $X: (\Omega, A, P) \longrightarrow R$

$$w \rightsquigarrow X(w)$$

tal que:

$$\forall\, r \in R,\ \{w \in R \,/\, X(w) \leq r\} \in A \text{ (é un suceso)} \qquad [*]$$

[*] CONDICIÓN: Para todo número real, o conxunto das antiimaxes de si mesmo e de todos os reais menores a el é un suceso.

• **VARIÁBEL ALEATÓRIA DISCRETA** é unha v.a. que toma un número de valores <u>finito ou infinito numerábel</u>: corresponden a experimentos nos que <u>se conta o número de veces</u> que acontece un suceso.

Unha v.a. discreta toma os valores $x_1, ..., x_n$ con $x_1<x_2<...<x_n$

▪ <u>**Función de Probabilidade ou Masa Px**</u> asociada a unha v.a. <u>discreta</u>

$$P_x: R \longrightarrow [0,1] \subset R \quad / \quad P_x(X) = \begin{cases} P_i = P(w/x(w)=x_i) & \text{se } x=x_i \quad i=1...n \\ 0 & \text{se } x \neq x_i \end{cases}$$

indica a probabilidade de que suceda un suceso elemental.

[O eixo OX só ten pontos “salteados” x_i cuia probabilidade P_x (altura respeito de OY) pode ser maior que 0, e como máximo até 1:]

Exemplo:

Representar a función de probabilidade ou masa da variábel aleatória discreta x ≡ *resultado de lanzar un dado*

$x(\omega) \in \{1,2,3,4,5,6\}$

```
      Px
      |
 1/6  +  .  .  .  .  .  .
      |
      +--o--o--o--o--o--o--  xi
         1  2  3  4  5  6
```

Obsérve-se que:

$$\mathbf{0 \leq P_i \leq 1}$$

$$\mathbf{\sum^{n} P_i = 1}$$

i=1

▪ **Función de Distribución F** asociada a unha v.a. discreta

F: R $\longrightarrow$ R
$x_0 \rightsquigarrow F(x_0)=P(x\le x_0)=\sum_{x_i\le x_0} P_i$

Propriedades:

$0\le F(x)\le 1$

$F(-\infty)=0$ $F(+\infty)=1$

F é crecente e contínua pola direita.

P(a<x≤b)=F(b)-F(a)

A función de distribución dunha v.a. discreta é unha función F(x) escalonada, crecente, con saltos nos x_i ($x_1...x_n$) iguais á probabilidade en cada x_i ($P_1...P_n$)

Propriedades dunha función de distribución F dunha variábel discreta x

1) F é non decrecente
2) F é contínua pola dereita $\lim_{x\to x_0^+} F(x)=F(x_0)$
3) $\lim_{x\to-\infty} F(x)=0$ (entre 0 e 1)
$\lim_{x\to+\infty} F(x)=1$

Relación entre unha función de distribución F dunha variábel discreta x e a función de probabilidade

1) $P(a<x\le b)=F(b)-F(a)=P(x\le b)-P(x\le a)$
2) $P(a\le x\le b)=F(b)-\lim(x\to a)\ F(x)$
3) $P(a<x<b)=\lim(x\to b)\ F(x)-F(a)$
4) $P(a\le x<b)=\lim(x\to b)\ F(x)-\lim(x\to a)\ F(x)$
5) $P(x=a)=F(a)-\lim(x\to a)\ F(x)$

Exemplo:
Representar a función de distribución da variábel aleatória discreta x ≡ *resultado de lanzar un dado*

```
                                  F(x)
                                   |
                            6/6    +           [→
                                   |
x(ω)∈{1,2,3,4,5,6}          5/6    +         [→
                                   |
F(1)=P(x≤1)=1/6             4/6    +       [→
F(2)=P(x≤2)=2/6                    |
...                         3/6    +     [→
                                   |
                            2/6    +   [→
                                   |
                            1/6    + [→
                                   |
                                   +→+—+—+—+—+—+—  xi
```

1 2 3 4 5 6

▪ Exercicio ***O número meio de accidentes dunha autopista ao ano é de 15. Calcular***

a) a probabilidade de dous accidentes nun mes
b) a probabilidade de tres accidentes nun ano
c) *a probabilidade de mais de dous accidentes nun trimestre*

sucesos discretos sobre soporte contínuo (tempo) ⇒ proceso de Poisson, distribución de Poisson ⇒

$$P(x=i)=\frac{e^{-\lambda}\cdot\lambda^{i}}{i!}$$

$E(x)=\lambda=15$
$Var(x)=\lambda=15$

a)

para calcular a probabilidade de que haxa 2 accidentes nun mes non podemos supoñer que haberá 24 nun ano, senón adaptar $\lambda=15$ á medida de meses: $E(x)=Var(x)=15/12$ sendo $x_{por\ mes}$

$$P(x_{por\ mes}=2)=\frac{e^{-(15/12)}\cdot(15/12)^{2}}{2!}=0.2238318\approx 0.22$$

táboa [Scheaffer/McClave p.645] é acumulativa
$\Rightarrow P(x=\mathbf{2},\lambda=1.2)=P(x\le\mathbf{2},\lambda=1.2)-P(x\le\mathbf{1},\lambda=1.2)=0.879-0.663=0.216\approx0.22$
táboa [fotocopias p.602] é acumulativa
$\Rightarrow P(x=\mathbf{2},\lambda=1.2)=P(x\le\mathbf{2},\lambda=1.2)-P(x\le\mathbf{1},\lambda=1.2)=0.8795-0.6626=0.2169\approx0.22$

b)

$$P(x=3)=\frac{e^{-15}\cdot15^{3}}{3!}=0.000172\approx 0.000$$

táboa [Scheaffer/McClave p.648] é acumulativa
$\Rightarrow P(x=\mathbf{3},\lambda=15)=P(x\le\mathbf{3},\lambda=15)-P(x\le\mathbf{2},\lambda=15)=0.000-0.000=0.000$
táboa [fotocopias p.609] é acumulativa
$\Rightarrow P(x=\mathbf{3},\lambda=15)=P(x\le\mathbf{3},\lambda=15)-P(x\le\mathbf{2},\lambda=15)=0.0002-0.0000=0.0002\approx0.000$

c)

$\lambda_{ano}=15 \Rightarrow \lambda_{trimestre}=15/(12/3)=15/4=3.75$

considerando x trimestral:

$P(x>2)=1-P(x\le2)=1-0.2854=0.7146$
[fotocopias p.603 x=2, $\lambda=3.7$]

- Exercicio *Consideremos o experimento aleatório do lanzamento de duas moedas.*
  - *a) definir a v.a. do número de caras obtidas*
  - *b) calcular a función de masa de probabilidade e a sua función de distribución*

a)

x: Ω={cc,cx,xc,xx} ⟶ R

cc ⇝ 2 caras
cx ⇝ 1 caras
xc ⇝ 1 caras
xx ⇝ 0 caras

b)

función de masa de probabilidade

P(x=2)=1/4
P(x=1)=2/4
P(x=0)=1/4

función de distribución

$$F: R \longrightarrow R \qquad t \leadsto F(t)=P(x\le t)=\begin{cases} 0 \text{ se } t<0 \\ ¼ \text{ se } 0\le t<1 \\ ¾ \text{ se } 1\le t<2 \\ 1 \text{ se } 2\le t \end{cases}$$

- Exercicio *Dada a seguinte táboa, calcular a función de distribución e as probabilidades P(x£4.5) e P(x>2)*

x	0	1	2	3	4	5
P(x)	1/10	3/10	4/10	1/10	1/20	1/20

F:R⟶R
t⇝F(t)=P(x≤t)=

0 se t<0
1/10 se 0≤t<1
4/10 se 1≤t<2
8/10 se 2≤t<3
9/10 se 3≤t<4
19/20 se 4≤t<5
1 se 5≤t

P(x≤4.5)=19/20
P(x>2)=1-P(x≤2)=1-8/10=2/10

• **VARIÁBEL ALEATÓRIA CONTÍNUA** é unha v.a. que toma valores nun intervalo contínuo, sexa finito ou infinito. Na prática non se pode coñecer o valor exacto dunha v.a. contínua, xa que medi-la equival a clasifica-la nun intervalo mais ou menos fino.

▪ **Función de Densidade f(x)** asociada a unha v.a. contínua

é o límite dos histogramas asociados ao experimento aleatório cando o tamaño da mostra tende a infinito e a base do histograma tende a cero $f(x):R \to R$

Redúce-se a base dos rectángulos e dun mesmo intervalo tóman-se mais valores, daquela a sucesión de rectángulos convírte-se nunha curva cada vez mais suave.

$f(x) \geq 0 \qquad \int f(x)dx = 1$ a área baixo a curva é 1

A integración da función de densidade indica a probabilidade de calquera suceso:

$$\mathbf{P(x \leq x_0) = \int_{-\infty}^{x_0} f(x)dx} \qquad \mathbf{P(a \leq x \leq b) = \int_a^b f(x)dx}$$

A probabilidade de que a v.a. tome valores nun intervalo ven dada pola área que f(x) encerra no intervalo.

$P(x=x_0)=0$ porque x é unha v.a. contínua, pero

$P(x_0-\Delta x \leq x \leq x_0+\Delta x) \approx$ base x altura $=2\Delta x \cdot f(x_0)$

polo tanto f(x) mide a densidade de probabilidade nun ponto.

A función de densidade permite calcular dunha forma sinxela as probabilidades de calquera suceso asociado aunha v.a. contínua.

▪ Exercicio *Para unha variábel aleatória contínua x con función de densidade f(x)=-x/2+1 calcular a probabilidade de que valla menos de 1.*

$$P(x<1) = \int_{-\infty}^{1} \frac{-x}{2}+1\,dx = \int_{0}^{1} \frac{-x}{2}+1\,dx = [-1/4 \cdot x^2+x]_0^1 = ¾ = 75\%$$

▪ Exercicio *Para unha variábel aleatória contínua x con función de densidade $f(x)=2e^{-2x}$ aproximar a probabilidade de que valla 3*

x v.a. contínua $\Rightarrow$ P(3)=0

$P(x_0-\Delta x \leq x \leq x_0+\Delta x) \approx 2\Delta x \cdot f(x_0) \quad \underset{\Delta x=0.0001}{\overset{\uparrow}{=}} 2\Delta x \cdot f(3) = 0.0002 \cdot 2 \cdot e^{-6}$ = 4 entre dez millóns

para $\Delta x=0.1$, $P(2.9 \leq x \leq 3.1) \approx 1$ entre mil

▪ Exercicio *Estudar a validez para unha v.a. contínua x das seguintes funcións como funcións de densidade...*

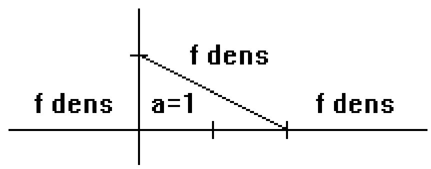
a) f(x)=-x/2+1 con x Î (0,2)
b) f(x)=1/b-a con x Î (a,b)
c) f(x)=le^-lx con x³0

As duas condicións que compre unha función de densidade é que a función sexa sempre positiva ou 0 e que a sua integral sexa 1.

a)

$$\int_0^2 \frac{-x}{2}+1\,dx = [-1/4\cdot x^2+x]_0^2 = -1+2 = 1$$

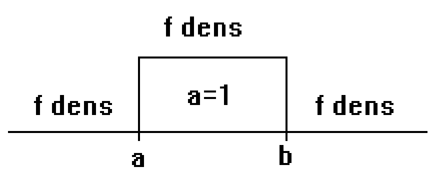
$f(x)\geq 0$



b)

$$\int_a^b \frac{1}{b-a}\,dx = [x/(b-a)]_a^b = 1$$

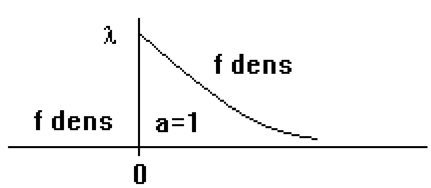
$f(x)\geq 0$



c)

$$\int_0^{+\infty} \lambda e^{-\lambda x}\,dx = [-e^{-\lambda x}]_0^{+\infty} = -e^{-\infty}+1 = -1/e^{\infty}+1 = 1$$

$f(x)\geq 0$



- **<u>Función de Distribución F</u>** asociada a unha v.a. contínua

$$F: \mathbb{R} \longrightarrow \mathbb{R}$$
$$x_0 \rightsquigarrow F(x_0)=P(x\leq x_0)=\int_{-\infty}^{x_0} f(x)dx$$

entón:

$$\left.\frac{dF(x)}{dx}\right|_{x=x_0} = f(x_0)$$

Propriedades:

$$0\leq F(x)\leq 1$$
$$F(-\infty)=0 \qquad F(+\infty)=1$$

A función de distribución dunha v.a. contínua é unha curva crecente e contínua

$P(a<x<b)=F(b)-F(a)$

Resume

v.a.discreta	función de probabilidade ou masa P_x $0\leq P_i\leq 1$ $\sum_{i=1}^{n} P_i=1$	función de distribución F $F(x_0)=P(x\leq x_0)=\sum_{xi\leq x0} P_i$ $0\leq F(x)\leq 1$ $F(-\infty)=0$ $F(+\infty)=1$
	$P(a<x\leq b)=F(b)-F(a)$	
v.a.contínua	función de densidade **f** $\mathbf{P}(x\leq x_0) = \int_{-\infty}^{x_0} \mathbf{f}(x)dx$ $\mathbf{P}(a\leq x\leq x_0) = \int_a^b \mathbf{f}(x)dx$ $P(x_0-\Delta x\leq x\leq x_0+\Delta x)\approx 2\Delta x\cdot f(x_0)$	función de distribución **F** $=\mathbf{F}(x_0)$ $0\leq F(x)\leq 1$ $F(-\infty)=0$ $F(+\infty)=1$
	$P(a<x<b)=F(b)-F(a)$	

▪ <u>Exercicio</u> *Estudar a función de distribución se a función de densidade é...*

a) f(x)=-x/2+1 con x Î (0,2)
b) f(x)=1/b-a con x Î (a,b)
c) f(x)=le^{-lx} con x³0

a)

$$F(x)=P(t\le x)=\int_{-\infty}^{x} \frac{t}{2}+1\,dt = [-1/4\cdot t^2+t]_{-\infty}^{x} = [-1/4\cdot t^2+t]_0^x = -1/4\cdot x^2+x$$

F contínua e crecente $F(-\infty)=0\le F(x)\le 1=F(+\infty)$

b)

$$F(x)=P(t\le x)= \int_{-\infty}^{x} \frac{1}{b-a}\,dt= \int_{a}^{x} \frac{1}{b-a}\,dt=[t/(b-a)]_a^x= (x-a)/(b-a)$$

F contínua e crecente $F(-\infty)=0\le F(x)\le 1=F(+\infty)$

c)

$$F(x)=P(t\le x)= \int_{0}^{x} \lambda e^{-\lambda x}\,dt = [-e^{-\lambda t}]_0^x = -e^{-\lambda x}+1 = -1/e^{\lambda x} +1$$

F contínua e crecente $F(-\infty)=0\le F(x)\le 1=F(+\infty)$

▪ <u>Exercicio</u> *A variábel aleatória que representa a proporción de accidentes automovilísticos mortais de EEUU ten por función de densidade a f(x)=42x(1-x)^5 se 0£x£1; 0 se toma outro valor. Cal é a probabilidade de que non mais do 25% dos accidentes automovilísiticos sexan mortais?*

$$P(x\le 0.25)= \int_{0}^{0.25} 42x(1-x)^5dx = \int_{0}^{0.25} udv = [uv]_0^{0.25} - \int_{0}^{0.25} vdu =$$

$$u=42x \Rightarrow du=42dx$$
$$dv=(1-x)^5dx \Rightarrow v=(-1)(1-x)^6/6$$

$$=[42x(-1)(1-x)^6/6]_0^{0.25}-0.75^7+1 = 0.566 \approx 57\%$$

▪ <u>Exercicio</u> *A variábel aleatória x representa o intervalo de tempo entre duas chegadas consecutivas a unha tenda e a sua densidade está dada por f(x)=k·exp(-x/2) se x>0. Calcular a probabilidade de que un intervalo entre duas chegadas consecutivas se atope entre 2 e 6 minutos.*

$$P(2<x<6)= \int_{2}^{6} k\cdot e^{-x/2}dx= [k\cdot e^{-x/2}\cdot(-2)]_2^6 = ...$$

canto val k?

$$\int_0^{+\infty} f(x)dx = 1 = [k\cdot e^{-x/2}\cdot(-2)]_0^{+\infty} = -k(-2) = 1 \Rightarrow k=1/2$$

$$...=-e^{-3}+e^{-1}=0.318$$

CAMBIO DE VARIÁBEL NA FUNCIÓN DE DENSIDADE f

Sexa x unha v.a. contínua con función de densidade $f_x(x)$ e y unha v.a. contínua relacionada con x pola *relación biunívoca* y=g(x), entón a función de densidade de y é

$$\boxed{f_y(y) = f_x(x(y))\left|\frac{dx(y)}{dy}\right|}$$

Exemplo:

Sexa x a v.a. con función de densidade $f_x(x)=k\cdot exp(-x/4)$ se x>0. Calcular a densidade da variábel y=2+(4/3)x

y=2+(4/3)x é unha función bixectiva ⇒ relación biunívoca
$x(y)=(y-2)3/4$
$f_x(x(y))=k\cdot exp((6-3y)/16)$
$|dx(y)/dy|=0.75$
$f_y(y)=0.75\cdot k\cdot exp((6-3y)/16)$ se x>0 ⇒ se y>0
$$\int_0^{+\infty} k\cdot exp(-x/4)=1=-4k\cdot[exp(-x/4)]_0^{+\infty}=-4k(-1)=4k \Rightarrow k=0.25$$
$f_y(y)=3/16\cdot exp(6/16-3/16y)$ se y>0

Se a relación non é biunívoca determínan-se todos os $x/x=g^{-1}(y_0)$ e súman-se todos os elementos de probabilidade correspondentes a eles. Sexan $x_1, x_2, ... x_n = g^{-1}(y_0)$, entón:

$$\boxed{f_y(y) = \sum_{i=1}^{n} f_x(x_i(y))\left|\frac{dx_i(y)}{dy}\right|}$$

Exemplo:

Sexa x a v.a. con función de densidade $f_x(x)=1/\sqrt{2\pi}\cdot exp(-x^2/2)$ Calcular a densidade de $y/x=y^2$

$x(y)=\pm\sqrt{y}$ a función non é bixectiva
⇒ a relación non é biunívoca
$dx=\pm1/(2\sqrt{y})dy \Rightarrow dx/dy=\pm1/(2\sqrt{y})$
$f_y(y)=f_x(+\sqrt{y})\cdot1/(2\sqrt{y})+f_x(-\sqrt{y})\cdot1/(2\sqrt{y})=$
$=1/(2\sqrt{y})[1/\sqrt{(2\pi)}e^{-1/2y}+1/\sqrt{(2\pi)}e^{-1/2y}]=$
$=1/(2\sqrt{y})\cdot2/\sqrt{(2\pi)}e^{-1/2y}=1/\sqrt{(2\pi y)}\cdot e^{-1/2y}$ se $y\geq0$

MEDIDAS CARACTERÍSTICAS DUNHA VARIÁBEL ALEATÓRIA

Os modelos de distribución de probabilidade son representacións idealizadas de experimetnos aleatórios. Representan unha suavización da variábel estatística obtida dunha mostra do experimento.

As medidas características dunha v.a. son as medidas características dunha variábel estatística cambiando P_i por f_i

		v.a.**DISCRETA** con valores $x_1 \ldots x_n$ e f.prob. $p_1 \ldots p_n$	v.a.**CONTÍNUA** con f.dens. f(x)
	MOMENTO RESP. ORIXE DE ORDE k $\boldsymbol{\alpha_k}$	$\alpha_k = \sum_{i=1}^{n} x_i^k \cdot p_i$	$\alpha_k = \int x^k f(x)dx$
$E(a+bx)=$ $=a+b\cdot E(x)$	ESPERANZA MATEMÁTICA media ou $\boldsymbol{\alpha_1}$	$E(x)=\mu= \sum_{i=1}^{n} x_i \cdot p_i$	$E(x)=\mu=\int xf(x)dx$
	ESPERANZA MATEMÁTICA DE y=g(x)	$E(y)=\mu= \sum_{i=1}^{n} g(x_i)\cdot p_i$	$E(y)=\mu=\int g(x)f(x)dx$
	MOMENTO RESP. MEDIA DE ORDE k $\boldsymbol{\mu_k}$	$\mu_k= \sum_{i=1}^{n} (xi-\mu)^k \cdot p_i$	$\mu_k=\int (x-\mu)^k f(x)dx$
$\sigma^2=E(x^2)-Ex^2$ $Var(ax+b)=$ $=a^2\cdot Var(x)$	VARIANZA $\boldsymbol{\mu_2}$ $\sigma^2=var(x)$ σ=desv.típica	$\mu_2= \sum_{i=1}^{n} (xi-\mu)^2 \cdot p_i$	$\mu_2=\int (x-\mu)^2 f(x)dx$

Coeficiente de Asimetria $CA=\mu_3/\sigma^3$

Coeficiente de Apuntamento ou Curtose $CP=\mu_4/\sigma^4-3$

<u>Acotación de Tsebyshev</u>

(indica a bondade de σ como medida de dispersión respeito da media μ)

"Sexa x unha v.a. (discreta ou contínua) con media μ e desviación típica σ, e $k\in R^+$, entón <u>a proporción da distribución</u> comprendida entre <u>$\mu-k\sigma$ e $\mu+k\sigma$</u> é <u>ao menos de $1-1/k^2$</u>"

$$P(\mu-k\sigma\le x\le\mu+k\sigma)=P(|x-\mu|\le k\sigma)\ge 1-\frac{1}{k^2}$$

En xeral, Tsebyshev afirma $P(g(x)>c)\le \dfrac{E(g(x))}{c}$

Un caso particular de T con interés é $g(x)=(x-\mu)^2$ e $c=k^2\sigma^2$

$$P((x-\mu)^2>k^2\sigma^2) \leq \frac{E(x-\mu)^2}{k^2\sigma^2}$$

||

$P(|x-\mu|>k\sigma) \leq 1/k^2$ xa que $P(|x-\mu|\leq k\sigma) \geq 1-1/k^2$