Apunts de teoria de la probabilitat

ALEIX TORRES I CAMPS

ANNA DE MIER (ANNA.DE.MIER@UPC.EDU), GUILLEM PEREARNAU I SONIA PEREZ

1 Espais de probabilitat

1.1 Motivació

L'objectiu de la teoria de la probabilitat és trobar models per a fenònems que depenen de l'atzar (no deterministes), cada realització d'un fenomen en direm experiment, del qual n'obtindrem un resultat. A més, tindrem els successos (observables) que son totes les preguntes raonables que ens podem fer.

1.2 Experiments i probabilitat

Definició 1. Un experiment és un parell $(\Omega, \mathscr{A})$ on Ω és un conjunt i $\mathscr{A} \subset \mathscr{P}(\Omega)$ tal que:

1. $\emptyset \in \mathscr{A}$
2. $A \in \mathscr{A} \implies A^c \in \mathscr{A}$
3. Si $\{A_n\}_{n\geq 1}$ és una col·lecció numerables d'elements de $\mathscr{A} \implies \bigcup_{n\geq 1} A_n \in \mathscr{A}$

Exemple 1. Uns quants exemples...

Volem una funció que assigni probabilitats de successos, és a dir, $P : \mathscr{A} \to \mathbb{R}$. Llavors definim:

Definició 2. Un espai de probabilitat és una terna $(\Omega, \mathscr{A}, P)$ on:

1. $(\Omega, \mathscr{A})$ és un experiment.
2. $P : \mathscr{A} \to \mathbb{R}$ tal que: $P(\emptyset) = 0$, $P(A) \geq 0$, $\forall A \in \mathscr{A}$. Si $\{A_n\}_{n\geq 1}$ és una col·lecció de successos dos a dos dijunts $\implies P(\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} P(A_n)$.
3. $P(\Omega) = 1$.

Per tant, la probabilitat és una mesura a $(\Omega, \mathscr{A})$ normalitzada a 1. A P se l'anomena funció de probabilitat.

Exemple 2. Espia discret, si Ω és numerable i $\mathscr{A} = \mathscr{P}(\Omega)$. Si $\Omega = \{w_1, w_2, w_3, \ldots\}$ prenem $\sum_{i\geq 1} p_i = 1$ (amb $p_i \geq 0$) i definim $\mathscr{P}(A) = \sum_{w_i \in A} P(\{w_i\})$, alleugerint la notació podem fer servir $P(\{w_i\}) = P(w_i) = p_i$.

Exemple 3. Espai clàssic, és un éspai discret amb $|\Omega| = N$ i $p_i = 1/N$. Çassos favorables entre cassos possibles": $P(A) = \frac{|A|}{N}$.

Exemple 4. Espais clàssics amb monedes o daus, tot ben repartit.

Exemple 5. Durada d'un mòbil? $\Omega = (0, \infty)$ o bé, $(0, L]$. Si $\mathscr{A} = \mathscr{P}(\Omega)$ sabem que no podem assignar-hi una mesura. Però com ens interessen els intervals com (a, b), agafe, la σ-àlgebra que conté tots els intervals (oberts). Aquí apareixen els burelians $\mathcal{B} = \sigma(I)$ i podem agafar la mesura de Lebesque a $\mathbb{R}$. En resum, $\Omega = (0, L)$, $\mathrm{B} = \sigma(I)$ i $P(B) = \frac{\mu(B)}{L}$. On $\mu(B)$ és la seva mesura de Lebesgue. Tot i així, no és realistic perquè és massa uniforme.

Proposició 3. *Propietats d'espais de probabilitat. Sigui* $(\Omega, \mathscr{A}, P)$*:*

1. *Per $r \geq 2$, si $A_1, \dots, A_r \in \mathscr{A}$, $A_i \bigcap A_j = \emptyset$ si $i \neq j$ llavors $P(\bigcup_{i=0}^r a_i) = \sum_{i=1}^r P(A_i)$*
2. *Si $A, B \in \mathscr{A}$ i $A \subset B \implies P(B \setminus A) = P(B) - P(A)$ i $P(A) \leq P(B)$.*
3. *$P(A^c) = 1 - P(A)$, $\forall A \in \mathscr{A}$*
4. *(Desigualtat de Boole) Si $A_1, \dots, A_r \in \mathscr{A} \implies P(\bigcup_{i=1}^r A_i) \leq P(A_1) + \cdots + P(A_r)$*

Demostració.

1. En els cassos finits, per $r < k$, cal agafar $A_k = \emptyset$, ja que així, com que $P(\emptyset) = 0$, $P(\bigcup_{1 \leq n \leq r}) = P(\bigcup_{1 \leq n} A_n) = \sum_{1 \leq n} P(A_n) = \sum_{1 \leq n \leq r} P(A_n)$.
2. Primer de tot $B \setminus A \in \mathscr{A}$, ja que $B \setminus A = (B^c \bigcup A)^c \in \mathscr{A}$. Després, reordenant el fet que $P(A) + P(B \setminus A) = P(B)$, ens queda el que volíem. Com les propietats són positives, la desigualtat es demostra automàticament.
3. De $P(A) + P(A^c) = P(\Omega) = 1$ obtenim l'expressió de l'enunciat.
4. Ho anem a fer per inducció sobre r. Clarament per $r = 1$ és cert, suposem que ho és per $r - 1$, anem a veure-ho per a un r arbitrari. Sigui $B = (\bigcup_{i=1}^{r-1} A_i) \bigcap A_r$, llavors $P(\bigcup_{i=1}^r A_i) = P((\bigcup_{i=1}^{r-1} A_i) \bigcup A_r) = P((\bigcup_{i=1}^{r-1} A_i) \bigcup (A_r - B_r)) = P((\bigcup_{i=1}^{r-1} A_i)) + P(A_r - B_r) =$ [per hipòtesi i per 2] $\leq P(A_1) + \cdots + P(A_{r-1}) + P(A_r)$ que és la desigualtat de Boole.

□

Proposició 4. *Successions monòtones:*
Si $A \in \mathscr{A}$, $i \geq 1$ i $A_1 \subset A_2 \subset A_3 \subset \cdots$, aleshores $P(\bigcup_{i \geq 1} A_i) = \lim_{i \to \infty} P(A_i)$.
Si $A \in \mathscr{A}$, $i \geq 1$ i $A_1 \supset A_2 \supset A_3 \supset \cdots$, aleshores $P(\bigcap_{i \geq 1} A_i) = \lim_{i \to \infty} P(A_i)$.

Demostració. Fem $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_2$, ... Aleshores, $B_i \bigcap B_j = \emptyset$ per $i \neq j$ i $A_i = \bigcup_{j=1}^i B_i$. Per tant:

$$P(\bigcup_{n \geq 1} A_n) = P(\bigcup_{n \geq 1} B_n) = \sum_{n \geq 1} P(B_i) = \lim_{N \to \infty} \sum_{n=1}^N P(B_n) = \lim_{N \to \infty} P(A_N)$$

L'altre és demostra passant al complementari. □

Teorema 5. *Siguin $A_1, \dots, A_r \in \mathscr{A}$. Per $I \subset [r]$, posem: $A_I = \bigcap_{i \in I} A_i$ i $S_k = \sum_{I \subset [r] \ |I| = k} P(A_I)$. Aleshores,*

$$P(\bigcup_{i=1}^r A_i) = \sum_{k=1}^r (-1)^{k+1} S_k$$

Demostració. Per inducció, el cassos $r = 1, 2$ són fàcils. Aleshores, pel cas inducctiu fa falta el cas $r - 1$ i el cas 2. □

Proposició 6. *Desigualtats de Bonferroni. Sigui $M_T = \sum_{k=1}^T (-1)^{k+1} S_k$. Aleshores, si:*
1. T és senar $\implies P(\bigcup_{i=1}^r A_i) \leq M_T$.
2. T és parell $\implies P(\bigcup_{i=1}^r A_i) \geq M_T$.

Demostració. La demostració per inducció és semblant a l'anterior. □

1.3 La probabilitat condicionada

Sigui $(\Omega, \mathscr{A}, P)$ un espai de probabilitat. Prenem $B \in \mathscr{A}$ amb $P(B) > 0$. Volem recalcular la probabilitat P dels successos sabent que ha passat B.

Definició 7. Si $B \in \mathscr{A}$ amb $P(B) > 0$ i $A \in \mathscr{A}$, la probabilitat de A condicionada a B és

$$P(A|B) = \frac{P(A \bigcap B)}{P(B)}$$

Observació 8.

1. $P(A|B)$, a priori, pot ser major o menor a $P(A)$.
2. Fixat B, $P_B : \mathscr{A} \to \mathscr{R}$ definida com $P_B(A) = \frac{P(A \bigcap B)}{P(B)}$, dona una funció de probabilitat en $(\Omega, \mathscr{A})$. (També ho és en $(\Omega, \mathscr{A}_B = \{A \bigcap B : A \in \mathscr{A}\})$).

Sigui ara $B_1, \cdots, B_n$ una partició de Ω (amb $B_i \in \mathscr{A}$ i $P(B) > 0$). Llavors, la llei de les probabilitats totals és

$$P(A) = P(A \bigcap \Omega) = P(A \bigcap (\bigcup B_i)) = P(\bigcup (A \bigcap B_i)) = \sum_{i=1}^{n} P(A \bigcap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

Exemple 6. Ruïna del jugador (Huygens S.XVII). Sigui J un jugador que comença amb un capital de $k \geq 1$, el seu objectiu és arribar a $N \geq k$ i s'arruina si arriba a 0. En cada torn guanya 1 amb probabilitat 1/2 i perd 1 amb probabilitat 1/2.

Sigui $B =$"a la 1a jugada, +1ï $R_k =$"s'ha arruinat amb capital inicial k". Aleshores,

$$P(R_k) = P(R_k|B)P(B) + P(R_k|B^c)P(B^c) = \frac{1}{2}P(R_{k+1}) + \frac{1}{2}(R_{k-1})$$

Definint $p_k = P(R_k)$, aleshores, obtenim l'eqüació de recurrència: $p_k = \frac{1}{2}p_{k+1} + \frac{1}{2}p_{k-1}$ amb $p_0 = 1$ i $p_N = 0$.

Per resoldre'l, fem $\frac{1}{2}(p_k - p_{k-1}) = \frac{1}{2}(p_{k+1} - p_k)$, definim $b_k = p_k - p_{k-1}$ que per la própia recurrència es compleix que $p_k = p_{k-t} + tb_{k-t+1}$. I veien que la solució és $p_k = 1 - \frac{k}{N}$ que es pot comprovar per inducció.

Pregunta: Qui era $(\Omega, \mathscr{A}, P)$?

Proposem $\Omega = \{w_1, w_2, w_3, \ldots) : w_i \in \{0, 1\}\}$, a cada successió li associem un real (no és injectiva, però només es repeteix en alguns racionals). $P(A) = \mu(\phi(A))$ on μ és la mesura de Lebesgui a $[0, 1]$ i *phi* passa de $(w_1, w_2, \ldots)$ a $0.w_1 w_2 \ldots$, llavors $\mathscr{A}$ és el conjunt de conjunts que van a parar a borelians per *phi*)

Teorema 9. *Bayes.* $P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum P(A|B_i)P(B_i)}$

Demostració. Bé del fet que $P(A \bigcap B_i) = P(B_i \bigcap A)$ i escrivint la definició de probabilitat condicionada de dues maneres diferents. I després, utilitzant el lemma de les probabilitats totals. □

1.4 Independència

R.Durret "Aquí acaba la teoria de la mesura i comença la probaibilitat".

Definició 10. En $(\Omega, \mathscr{A}, P)$", dos scuccessos $A, B \in \mathscr{A}$ són independents si $P(A \bigcap B) = P(A) \cdot P(B)$.

Observació 11. *Si $P(B) > 0$, $P(A|B) = \frac{P(A \bigcup B}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$.*

Exercici: Proveu que el conjunt buit i el total son independents amb qualsevol altre succés.

Demostració. La probabilitat del buit sempre és 0 i si intersequem quelcom amb el buit sempre el dona el buit, per tant, sempre es compleix que $0 = P(\emptyset)P(A) = P(\emptyset \bigcap A) = 0$.

Com que la probabilitat del total és 1, succeix el següent: $P(A)P(\Omega) = P(A) = P(A \bigcap \Omega)$. □

Exercici: Proveu que si A i B són independents $\implies$ que amb o sense complementaris també ho són.

Demostració. Només cal comprovar que A^c i B són independents, perquè veure que A i B^c són independents és el raonament simétric i per obtenir que A^c i B^c cal fer el mateix raonament dues vegades.

Simplement fem $P(A^c)P(B) = (1 - P(A))P(B) = P(B) - P(A)P(B) = P(B) - P(A \bigcap B) = P(A^c \bigcap B)$. □

Definició 12. Si $\{A_i\}_{i \in I}$ és una col·lecció de successis, són independts si $\forall J \subseteq I,\ |J| < \infty,\ P(\bigcap_{j \in J} A_j) = \prod_{i \in K} P(A_j)$.

1.5 Espais productes

Tenim $(\Omega_1, \mathscr{A}_1, P_1)$ i $(\Omega_2, \mathscr{A}_2, P_2)$ dos espais de probabilitat. Volem un espai en $\Omega_1 \times \Omega_2$ tal que $P(A_1 \times A_2) = P_1(A_1)P_2(A_2)$ ($\forall A_i \in \mathscr{A}_i$ per $i = 1, 2$). La teoria de la mesura ens diu que es pot... I la σ-àlgebra. Com volem garantir que hi hagi $\mathscr{A}_1 \times \mathscr{A}_2 = \{A_1 \times A_2 : A_i \in \mathscr{A}_i, i = 1, 2\}$. Tot i així, això no tñe perquè ser un σ-àlgebra. Llavors agafem la més petita que ho contingui. Com a $\mathscr{A}$ agafem $\sigma(\mathscr{A}_1 \times \mathscr{A}_2)$ és a dir, la generada. La probabilitat ? Volem que $P(A_1 \times A_2) = P(A_1)P(A_2)$.

Definició 13. Una àlgebra i una premesura són una σ-àlgebra i un mesura (respectivament, però només garanteix la unió finita i la suma finita.

Teorema 14. *Teorema d'extensió. Sigui p_0 una premesura en una álgebra $\mathscr{A}_0$. Aleshores existeixen una σ-àlgebra $\mathscr{A}^* \subset \sigma(\mathscr{A}_0)$ i una mesura p^* tal que p^* coincideix amb p_0 en $\mathscr{A}_0$. A més, si p_0 és finita, p és única.*

Com s'aplica?

$\mathscr{A}_0 = \{A \subset \Omega_1 \times \Omega_2|$ on A és unió finita d'elements de $\mathscr{A}_1 \times A_2\}$. (Caldria comprovar que és àlgebra i la unió es pot prendre disjunta i finita).

Exemple 7. Problema de l'agulla del comte Buffon (s. XVIII). I es pregunta, quina és la probabilitat que l'agulla talli alguna de les línies?

Sigui d la distància del punt mig a la recta més propera i sigui α l'angle amb la vertical. Aleshores, la longitud del catet $\frac{l}{2} \sin \alpha$. Tallarà si $d \leq \frac{l}{2} \sin \alpha$.

Així que com a espai mostral podem agafar $\Omega = [0, \pi) \times [0, \frac{L}{2}]$ la primera correspon a α i la segona a d. Amb la mesura de Lebesgue (volem uniformitat). Que finalment, calculem l'àrea sota la curva $l/2 \sin \alpha$ i dividim pel total, dona $\frac{2l}{L\pi}$.

1.6 Lemes de Borel-Cantelli

Sigui $(\Omega, \mathscr{A}, P)$ un espai de probabilitat i siguin $\{A_n\}_{n \geq 1}$ una col·lecció numerable de succesos.

Definició 15. Sigui $\limsup A_n = \bigcap_{n \geq 1} \bigcup_{k \geq n} A_k$ i $\liminf A_n = \bigcup_{n \geq 1} \bigcap_{k \geq n} A_k$. Són d'$\mathscr{A}$ perquè fem interseccions i unions numerables.

Nota 8. *Si $w \in \Omega$ i $w \in \limsup A_n \iff \forall n \geq 1,\ w \in \bigcup_{k \geq n} A_k \iff \forall n \geq 1 \exists k \geq n,\ w \in A_k \iff w$ pertany a infifnits dels A_n.*

Nota 9. *Fent el mateix amb $\liminf$, tenim $w \in \liminf A_n \iff w$ pertany a tots els A_n a partir d'algun en endavant.*

Nota 10. *Queda clar que* $\liminf A_n \subseteq \limsup A_n$.

Lema 16. *(Borel-Cantelli 1) Si $\sum_{n \geq 1} P(A_n) < \infty$, aleshores $P(\limsup A_n) = 0$.*

Demostració. Volem fitar $P(\limsup A_n) = P(\bigcap_{n\geq 1}\bigcup_{k\geq n} A_k)$ =[Succesions monòtones]= $\lim_{n\to\infty} P(\bigcup_{k\geq n} A_k) \leq \lim_{n\to\infty}\sum_{k\geq n} P(A_k) = 0$, perquè la cua d'una série convergent tendeix a 0. □

Pregunta natural: Què passa si $\sum_{n\geq 1} P(A_n) = \infty$?

Lema 17. *(Borel-Cantelli 2) Si $\sum_{n\geq 1} P(A_n) = \infty$ i els $\{A_n\}_{n\geq 1}$ són independents (dos a dos és suficient) $\implies P(\limsup A_n) = 1$.*

Demostració. $P((\limsup A_n)^c) = P(\bigcup_{n\geq 1}\bigcap_{k\geq n} A_k^c)$. Fixem-nos que això és el mateix que $\liminf A_k^c$ i que els successos estan encaixats $\bigcap_{k\geq 1} A_k^n \subseteq \bigcap_{k\geq 2} A_k^c \subseteq \cdots$. Llavors, per la proposició de successos encaixats tenim que la igualtat anterior és igual a $= \lim_{n\to\infty} P(\bigcap_{k\geq n} A_k^c)$, fent servir de nou la proposició dels successos encaixats tenim que $P(\bigcap_{k\geq n} A_k^c) = \lim_{r\to\infty} P(\bigcap_{k=n}^r A_k^c) = [ind.] = \lim_{r\to\infty}\prod_{k=n}^r (1 - P(A_k)) \leq \lim_{r\to\infty}\prod_{k=n}^r e^{-P(A_k)}$ (Utilitzem que $1 - x \leq e^{-x}$, per $x \geq 0$). Finalment, l'últim terme és $= \lim_{r\to\infty} e^{-\sum_{k=n}^r P(A_k)} = 0$. Per tant, com que cada un dels $P(\bigcup_{k\geq n} A_k^c) = 0$, $\lim_{n\to\infty} P(\bigcap_{k\geq n} A_k^c) = 0$. □

2 Variables aleatòries

2.1 Definició i distribució

Definició 18. L'aplicació $X : \Omega \to \mathbb{R}$ és una variable aleatória si $\forall B \in \mathscr{B}(\mathbb{R})$, es té que $X^{-1}(B) \in \mathscr{A}$ (on $\mathscr{B}$ són els borelians).

Observació 19. *X v.a. $\iff$ X és mesurable respecte $(\mathbb{R}, \mathscr{B})$.*

Observació 20. *X v.a. $\iff$ $\forall a \in \mathbb{R}$, $X^{-1}((-\infty, a]) \in A$.*

Notació: Escriurem $P(X \in B) = P(X^{-1}(B)) = P(\{w \in \Omega : X(w) \in B\})$. Igualment $P(X \geq x), P(X = x), P(X > x), \ldots$

Definició 21. La *funció de distribució (acumulada)* d'una v.a. X és $F_X : \mathbb{R} \to [0, 1]$ que envia $x \mapsto P(X \leq x) = P(\{w \in \Omega : X(w) \leq x\})$.

Exemple 11. La *variable indicadora* del succés $A \in \mathscr{A}$ és $\mathbb{I}_A : \Omega \to \mathbb{R}$ que envia $w \in \Omega$ a 1 si $w \in A$ i 0 altrament. Llavors, la funció indicadora $F_{\mathbb{I}_A}$ és igual a $P(x \leq x_0) = 0$ fins arribar a $x_0 = 0$. A partir de llavors, entre $[0, 1)$ la funció de distribució acumulada és $1 - P(A)$ i, a partir de A és igual a 1.

Proposició 22. *Si X és v.a. F_x satisfà:*

i) *F_x és creixent.*

ii) $\lim_{x\to+\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$.

iii) *F_X és contínua per la dreta* ($\lim_{h\to 0^+} F_X(x + h) = F_X(x)$)

Demostració.

i) Si $x_1 \leq x_2$ $F_X(x_1) = P(X \leq x_1) = P(\{w \in \Omega : X(w) \leq x_1\}) \leq P(\{w \in \Omega(w) : X(w) \leq x_2\}) = F_X(x_2)$

ii) Quan fem $\lim_{x\to+\infty} F(x) = \lim_{x\to+\infty} P(X \leq x) = \lim_{x\to+\infty} P(\{w \in \Omega : X(w) \leq x\}) = P(\Omega) = 1$ i el mateix quan x tendeix a menys infinit, el conjunt tendeix a $\emptyset$, per tant, té probabilitat 0.

iii) Sigui $\{h_n\}_{n\geq 1}$ una successió decreixent tal que $h_n \to 0^+$, llavors: $\lim_{n\to\infty} F_X(x + h_n) = \lim_{n\to\infty} P(\{w \in \Omega : X(w) \leq x + h_n\}) = P(\bigcap_{n\geq 1}\{w \in \Omega : X(w) \leq x + h_n\}) = P(\{w \in \Omega : X(w) \leq x\}) = F_X(x)$.

□

Nota 12. *Si $F : \mathbb{R} \to [0, 1]$ tal que satisfaci i), ii) i iii), aleshores existeix $(\Omega, \mathscr{A}, P)$ i $X : \Omega \to \mathbf{R}$ v.a. tal que $F_X = F$.*

3 V.a Discretes

4 V.a Contínues

5 Funcions característiques i famílies exponencials

6 Convergència de variables aleatòries