Inferència Estadística Avançada

November 2024

Pau Soler Valadés
pau.soler.valades@gmail.com

Abstract

Aquests apunts son un intent d'organitzar les meves idees i de omplir tots els buits que tinc en aquest bonic camp que és l'estadística

Contents

1 **Tema 0: Bàsics i Prèvis** 3
1.1 Espai de Probabilitats 3
1.2 Variables Aleatòries 3
1.3 Llei d'una Variable Aleatòria 4
1.3.5 Variables Aleatòries Discretes 5
1.3.6 Variables Aleatòries Absolutament Contíunes 5
1.4 Densitats de transformacions de Variables Aleatòries 6
2 **Tema 1: Reducció de Dades** 7
2.1 Estimadors i Reducció de Dades 7
2.2 Versemblança 7
2.3 Estimadors de Màxima Versemblança 8
Bibliography 9

1 Tema 0: Bàsics i Prèvis

1.1 Espai de Probabilitats

Definition 1.1.1 (Espai de Probabilitat). Un espai de Probabilitat és una terna $(\Omega, \mathcal{A}, P)$ on

- Ω *espai mostral*: conté tots els possibles resultats de l'experiència aleatòria.
- $\mathcal{A}$ *esdeveniments possibles*: familia de parts d'Ω ($\mathcal{A} \subset \mathcal{P}(\Omega)$) que té estructura de σ-àlgebra, que vol dir que és una σ-àlgebra si compleix que
  1. $\Omega \in \mathcal{A}$
  2. $\mathcal{A}$ és estable per pas al complementari ($A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$)
  3. $\mathcal{A}$ es estabale per unions numberables ($\{A_n, n \geq 1\} \subset \mathcal{A} \Rightarrow \cup_{n\geq 1} A_n \in \mathcal{A}$)
- P probabilitat: assigna versemblança a un element d'$\mathcal{A}$. $P : \mathcal{A} \longrightarrow [0,1]$. P compleix:
  1. $P(\Omega) = 1$
  2. $P\big(\cup_{n\geq 1}^{\text{inf}} A_n\big) = \sum_{n=1}^{\text{inf}} P(A_n)$

Recordatori ràpid de les propietats de la probabilitat P:

1. $P(\emptyset) = 0$
2. σ-additivitat implica additivitat finita. És a dir, la probabilitat de que n esdeveniments suceeixin és la suma de les seves probabilitats si els esdeveniments son disjunts. Si aquests no fossin disjunts, la probabilitat seria major o igual, en comptes d'igual.
3. $\forall A \in \mathcal{A} P(A^c) = 1 - P(A)$
4. $A, B \in \mathcal{A}, A \subset B \Rightarrow P(A) \leq P(B)$
5. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$

I també de les altres dues grans eines de la probabilitat:

Definition 1.1.2 (Probabilitat Condicionada). La probabilitat de que $A \in \mathcal{A}$ estigui condicionada per $B \in \mathcal{A}$ és

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

I se'n deriva inmediatament el següent

Proposition 1.1.3 (Formula Probabilitats Compostes). Siguin $A_1, ..., A_n \in \mathcal{A} \mid P(A_1 \cap ... \cap A_n) > 0$. Aleshores,

$$P(A_1 \cap ... \cap A_n) = P(A_1)P(A_2|A_1)...P(A_n|A_1 \cap ... \cap A_{n-1})$$

Proposition 1.1.4 (Fórmula de Bayes). Siguin A, B dos esdeveniments de probabilitat no nuls, es compleix que

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

1.2 Variables Aleatòries

A nivell matemàtic és incòmode treballar amb $\omega \in \Omega$, ja que els elements son bastant arbitràris. Per a tenir quelcom més uniforme, tractem amb variables aleatòries.

Definition 1.2.1 (σ-àlgebra de Borel). Denotem amb $\mathcal{B}$ la σ-àlgebra de Borel, que és la generada pels conjunts oberts d'$\mathbb{R}$. És a dir, $\mathcal{B}$ és la σ-àlgebra de $\mathcal{P}(\mathbb{R})$ que conté tots els oberts d'$\mathbb{R}$ respecte la topologia euclidiana.

Definition 1.2.2 (Variable Aleatòria). Una **Variable aleatòria** és una aplicació $X : \Omega \longrightarrow \mathbb{R}$ que compleix la condició de *mesurabilitat*, que és la següent:

$$\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{A}$$

La condició de mesurabilitat de la variable aleatòria permet assignar un valor numèric a qualsevol $\omega \in \Omega$ de la següent manera:

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

I per tant $X : \Omega \longrightarrow \mathbb{R}$ és una variable aleatòria.

Proposition 1.2.3 (Propietats de les variables aleatòries). Les propietats de les variables aleatories són, sent X, Y variables aleatòries:
- $aX + bY, XY$ també son variables aleatòries
- $\max(X, Y), \min(X, Y)$ són variables aleatòries
- $\frac{X}{Y}$ és variable aleatòria si $0 \notin \Im(Y)$
- L'ínfim, el suprem, el límit de l'ínfim i del suprem son variables aleatòries donada una successió de variables aleatòries
- Tota variable aleatòria no negativa és el límit creixent d'una successió de variables aleatòries simples positives.

1.3 Llei d'una Variable Aleatòria

Si les variables aleatòries et permeten assignar valors numèrics a elements de l'espai mostral, les lleis d'una variable aleatòria et permeten assignar probabilitats a aquests esdeveniments.

Definition 1.3.1 (Llei d'una Variable Aleatòria). La llei d'una variable aleatòria X és la probabilitat sobre $(\mathbb{R}, \mathcal{B})$, que anomenarem Q definida de la següent manera:

$$\forall B \in \mathcal{B}, Q(B) = P(X^{-1}(B)) \mid X^{-1}(B) = \{\omega, X(\omega) \in B\}$$

per tant, $Q = P \circ X^{-1}$ i clarament $Q : \mathcal{B} \to [0, 1]$

Es demostra trivialment que Q és una probabilitat sobre $(\mathbb{R}, \mathcal{B})$.

Per tant, amb aquesta definició hem aconseguit passar de l'espai de probabilitat $(\Omega, \mathcal{A}, P)$ a l'espai de probabilitat numèric $(\mathbb{R}, \mathcal{B}, P \circ X^{-1}$. Aquests segons espais son els que, per defecte, es fan servir quan es parla de probabilitats.

Definition 1.3.2 (Funció Distribució). La funció de distribució associada a una variable aleatòria X és la funció $F : \mathbb{R} \longrightarrow [0, 1]$ definia per

$$F(x) = P \circ X^{-1}((-\infty, x])$$

Remark 1.3.3 (Inducció de Probabilitat a X). Ara mateix, tenim dos espais de probabilitat, el numèric i el tradicional. El que s'aconsegueix amb la llei i la composició de les funcions és operar amb el millor dels dos mons. És a dir, la *mesurabilitat* de les variables aleatòries ens garanteix que quan volguem calcular la probabilitat d'un element a la recta real (i per tant membre de la σ-àlbegra $\mathcal{B}$) aquest SEMPRE equivalgui a un element a l'espai mostral Ω.

Per tant **tota variable aleatòria X indueix una probabilitat a** $(\mathbb{R}, \mathcal{B})$, i aquesta probabilitat induïda és Q, la llei d'una variable aleatòria. És a dir, quan informalment volem calcular $P(B), B \in \mathcal{B}$ (sent P aquí la idea de "probabilitat") d'una variable aleatòria, en realitat estem fent

$$Q(B) = P \circ X^{-}1(B) = P(\{\omega, X(\omega) \in B\})$$

unint l'espai numèric de probabilitat amb l'espai de probabilitat normal. Això és el que volem dir amb que X indueix una probabilitat a $(\mathbb{R}, \mathcal{B})$.

Per acabar, quan parlem de funció de distribució, seria l'equivalent a la Llei de la variable aleatòria però a la recta real.

Proposition 1.3.4 (Propietats Funció Distribució). Sigui F la funció de distribució d'una variable aleatòria X. Es compleix que:

1. F és creixent
2. F és contínua per la dreta ($\lim_{x\to a^-} F(x) = F(a)$)
3. $\lim_{x\to\infty} F(x) = 1$ i $\lim_{x\to-\infty} F(x) = 0$

Concretament, tot i existir $\lim_{y\to x^-} F(y) \forall c \in \mathbb{R}$, el que no podem esperar es que $F(x^-) = F(x) \forall x \in \mathbb{R}$, és a dir, F té discontinuïtats de primera espècie, i al ser acotada, té un nombre finit d'aquestes discontinuïtats.

El que ens dona la funció de disctribució de variables aleatòries és la capacitat de calcular les probabilitats a través d'aquesta. Siguin F_X la funció de distribució de la variable aleatoria X i P_X la probabilitat induïda per aquesta. Aleshores:

$$F_{X(x)} = P_X((-\infty, x]) = P(X^{-1}(-\infty, x]))) = P(\{\omega, X(\omega) \in B\}) = P(X \leq x)$$

I des d'aquí, podem obtenir les següents probabilitats naturalment des de F:

- Probabilitat d'un obert d'$\mathbb{R}$: $P((a, b]) = P(a < X \leq b) = F_X(b) - F_X(a)$
- Probabilitat de la unió d'oberts d'$\mathbb{R}$: $P(\cup_{i=1}^{\infty} (a_i, b_i]) = \sum_{i=1}^{\infty} P((a_i, b_i])$
- Probabilitat d'un punt d'$\mathbb{R}$: $P(X = x) = F(x) - F(x^-)$

1.3.5 Variables Aleatòries Discretes

Aquí posarem tots els tipus de distribucions per repassar amb les seves lleis, però sento que aquesta gent té bastanta més pressa ara per ara.

1.3.6 Variables Aleatòries Absolutament Contíunes

Definition 1.3.7 (Densitat). Una funció f és una densitat si compleix:

1. $f \geq 0$
2. f és integrabe Riemann a $\mathbb{R}$
3. Es té que:

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

Aleshores, amb la densitat podem definir:

Definition 1.3.8 (Variable Aleatòria Abolutament Contínua). Una variable aleatòria X és absolutament contínua amb densitat f si la funció de distribució es pot escriure com:

$$F(x) = \int_{-\infty}^{x} f(y)dy$$

on f és una densitat.

I des de la definició, és inmediat que:

$$P(X \in I) = \int_a^b f(x)dx, I \in \{[a, b], (a, b), [a, b), (a, b]\}$$

No tota variable contínua és absolutament contíuna, però el revers sí que és cert.

1.4 Densitats de transformacions de Variables Aleatòries

Pàgina 71 pdf Marta Sanz, com es fa Això

2 Tema 1: Reducció de Dades

2.1 Estimadors i Reducció de Dades

Primer defnim notació bàsica per esclarir-nos. Donada aquesta frase:

$Y_1, ..., Y_n$ és una mostra, definida per $Y \sim f(\boldsymbol{y}, \theta)$

Notation 2.1.1.
- Y_i és una variable aleatòria representant la mostra observada. No són encara dades, sinó com serien aquestes generades.
- Y és un vector aleatòri, és a dir, una variable aleatòria que representa la distribució, és a dir, una mostra de la distribució.
- $Y \sim f(\boldsymbol{y}, \theta)$ Y segueix una distribució amb densitat $f(\boldsymbol{y}, \theta)$, i $\boldsymbol{y}$ són les observacions $(y_1, ..., y_n)$ (és a dir, dades conegudes d'avantmà) depenent del parametre θ.

Definition 2.1.2 (Estadístic). Un **Estadístic** és una funció de la mostra $Y = (Y_1, ..., Y_n)$ que no depen del parametre θ

Sobre θ, és un parametre que modifica la distribució de les dades. Si és conegut, sabem exactament com la població Y es comporta. En general però, no s'acostuma a saber, i és l'objectiu d'aquest curs *estimar* θ correctament.

Per estimar un parametre, hem de fer tres distincions conceptuals:

Definition 2.1.3 (Parametre a estimar). El Parametre a Estimar (o Estimand) és la quantitat que no coneixem i a la que volem donar-li valor, el que seria θ.

Definition 2.1.4 (Estimador). Un **Estimador** és una funció que, donada la mostra, ens dona el valor de θ.

Definition 2.1.5 (Estimació). La **estimació** d'un parametre és el valor que s'obté amb l'estimador corresponent.

És a dir, si $X = (X_1, ..., X_n) \sim N(\mu, \sigma^2)$ amb variància desconeguda i volem trobar la mitjana:
- Parametre a estimar: $\theta = \mu$
- Estimador: $\bar{X}$ és la funció que estima θ
- Estimand: El valor de $\bar{X} = \sum_{i=i}^{n} \frac{X_i}{n} = \mu$

Definition 2.1.6 (Estimador Puntual). Un estimador puntual $\hat{\theta}$ de θ és qualsevol funció mesurable de la mostra $(Y_1, ..., Y_n)$ que no depèn de θ. Concretament, qualsevol estadístic és un estimador puntual.

Per exemple, $\hat{\theta} = \hat{\theta}(Y_1, ..., Y_n) = \frac{1}{n} \sum_{i=1}^{n} Y_i$ és una variable aleatòria.

Definition 2.1.7 (Estimació Puntual). Una estimació puntual $\hat{\theta}$ de θ és el resultat numèric del procés d'estimació, basat en els valors observats $(y_1, ..., y_n)$.

Per exemple, $\hat{\theta} = \hat{\theta}(y_1, ..., y_n) = \frac{1}{n} \sum_{i=1}^{n} y_i$ és un valor real. La diferència clau és que:
- L'estimador puntual treballa amb variables aleatòries $(Y_1, ..., Y_n)$
- L'estimació puntual treballa amb valors observats $(y_1, ..., y_n)$

Tot i que ambós es denotin com $\hat{\theta}$, representen conceptes diferents: una funció (estimador puntual) vs. un nombre (estimació puntual)

2.2 Versemblança

Definition 2.2.1 (Versemblança). Donada una FMP (FDP) $f(\boldsymbol{y}|\theta)$ d'una mostra $Y = (Y_1, ..., Y_n)$ i les observacions $\boldsymbol{y} = (y_1, ..., y_n)$, la funció de versemblança de θ és

$$L(\theta|\boldsymbol{y}) = f(\boldsymbol{y}|\theta) = \prod_{i=1}^{n} f_Y(y_i \mid \theta)$$

Versemblança vs Probabilitat

Aquí l'exemple de Poisson on es fan la mateixa pregunta a la inversa però obtens la mateixa resposta.

La diferència entre la probabilitat i la versemblança és només el teu coneixement a priori.

- Quina és la probabilitat de que $Y = 3$ si $Y\sim$ Poiss (2)? És a dir, ja sabem que $\lambda = 2$, i volem trobar com de probable és la nostra observació.
- Quina és "la probabilitat" de que les meves dades segueixen una Possion amb parametre $\lambda = 2$ si la observació és $Y = 3$? No sabem la distribució, concretament és el que volem estimar.

La versemblança ens respon la pregunta que volem contestar quan busquem un estimador: tenint les dades $\boldsymbol{y}$ quin parametre θ fa més versemblants les observacions? Com podem saber quins estimadors *maximitzen la versemblança*?

2.3 Estimadors de Màxima Versemblança

Bibliography