Tarea 1

Sara Luz Valenzuela Camacho　Valeria Durán Rubio
Luis Alberto García Aguilar　Rosendo Ivan García Alba

2024-01-22

Regresión Avanzada

16 de enero de 2024

Tarea 1

Fecha de entrega: 22 de enero 2024

1. Leer capítulos 1 y 2 de las notas de Bernardo, Bioestadística, que pueden encontrar en las lecturas del curso. Hacer un resumen de no más de tres cuartillas.

Capítulo 1: Introducción

1.1 Alcance y objetivos del libro

i. Extraer las conclusiones pertinentes de un conjunto de datos con estructura sencilla
ii. Resolver problemas de decisión moderdamente complicados
iii. Evaluar e contenido de las conclusiones de tipo estadístico publicadas en la lieratura científica

1.2 Estadística y Teoría de la Desición

La teoría de la decisión no sólo proporcionauna metodología para resolver problemas de decidión concretos sino que proporciona además una base teórica sobre la que puede construirse una metodología general unificada que contiene, como casos partticulares, aquellas recetas clásicas cuya utilidad ha sido demostrada con el tiempo.

La conclusión fundamental de esta investigación es que existe una única forma razonable de tomar decisiones. Aunque nauralmente existe una cierta libertad en la elección de los principios básicos, el resultado es siempre el mismo.

En primer lugar es necesario determinar el conjunto de las decisiones posibles y de aquellos sucesos cuya ocurrencia pueda modificar las consecuencias de la decisión tomada.

En segundo lugar debe cuantificarse mediante, mediante probabilidades, la verosimilitud asociada por el decisor a la ocurrencia de tales sucesos.

En tercer lugar deben describirse detalladamente las posibles consecuencias de cada una de las decisiones, las peferencias del decisor entre ellas deben ser evaluadas y cuantificadas en térmnos de una magnitud común que recibe el nombre de utilidad.

Finalmente debe tomarse aquella decisión que, con base a las probabilidades calculadas, proporcione la máxima utilidad esperada.

1.3 Probabilidad

La idea de que la probabilidad de un suceso en unas condiciones determinadas no es más que una medida de verosimilitud asociada por el decisor a la ocurrencia de tal suceso en esas condiciones, caracteriza la escuela de pensamiento Bayesiana.

Los métodos estaísticos clásicos, por otra parte, limitan el concepto de probabilidad a aquellos sucesos en los que puede definirse frecuencias relativas, lo que reduce seriamente su utilidad práctica, y no dispnen de una base axiomática en la que fundamentarse, por lo que frecuentemente dan lugar a resultados contradictorios.

1.4 Inferencia Estadística

La información inicial descrita por probabilidades iniciales, es combinada con los datos mediante el Teorema de Bayes para obtener las probabilidades finales, que describen la información total de que se dispone. Este proceso, que denominamos proceso de aprendizaje, conatituye la escencia de la metodología Bayesiana, cuyo nombre hace referencia precisamente al uso repetido del teorema de Bayes.

1.5 Problemas específicos de decisión

El estudio de las decisines sucesivas no plantea problemas nuevos; bastará resolver consecutivamente todos los problemas de decisión que integran la cadena, empezando por la decisión que debe ser tomada en el último lugar. En particular, el problema del diseño de experimentos puede ser planteado como una cadena de decisiones consecutivas. Se trata, en efecto, de elegir el experimento más adecuado para, una vez realizado, tomar la decisión que resulte más apropiadad a la vista de la información proporcionada por el experimento elegido

Capítulo 2: Fundamentos de la estadística y de la teoría de decisión

2.1 Estructura de un problema de decisión

El primer paso para resolver el problema de decisión es, pues, elaborar el conjunto de las posibles alternativas, al que llamaremos espacio de decisiones y denotaremos por D. La elección de uno de los elementos de D excluye la elección de cualquier otro elemento.

La principal dificultad con que uno se encuentra al plantearse un problema de decisión consiste en la falta de información sobre lo que sucederá según se actúe de una u otra manera. El problema general de decisión se plantea, pues, en ambiente de incertdumbre.

Una vez determinado el espacio de decisiones, habrá que considerar para cada una de las decisiones posibles el conjunto de sucesos inciertos que determinan sus eventuales consecuencias.

En resumen, nuestro modelo de problema de decisión exige la especificación del espacio $D = \{d_1, d_2, ..., d_k\}$ de las decisiones alternativas y de los conjutnos $\Theta_i = \{\theta_{i1}, \theta_{i2}, ..., \theta_{im_i}\}$ de sucesos inciertos mutuamente excluyentes que condicionan las consecuencias de cada una de las decisiones. El problema de decisión consiste entonces en elegir una decisión d_i del conjunto D sin saber cual de los sucesos θ_{ij} de Θ_i tendrá lugar.

2.2 Solución intuitiva a un problema de decisión

La probabilidad de un suceso A en una situación H, que representaremos como $p(A|H)$, será una medida sobre una escala $[0, 1]$ de la verosimilitud que se concede al suceso A en las condiciones descritas por H, esto es, una medida del grado de creencia en A que nos sugiere la información contenida en H. La probabilidad asignada a un suceso es siembre condicional a la informción que se psee sobre él: no existen probabilidades absolutas. Sin embargo, con objeto de simplificar la notación, la condición H será omitida cuando ello no dé lugar a confusión, escribiéndose $p(A)$ en lugar de $p(A|H)$.

Obviamente, el decisor tendrá sus preferencias entre las distintas consecuencias. En principio tales consecuencias podrían ser cuantificadas asignando a cada una de las consecuencias c_{ij} un número $u(c_{ij})$ que midiese la utilidad de que cada una del ellas tuviese para el decisor. La función de utilidad así construida mide las preferencias del decisor entre las posibles consecuencias de su decisión, de forma parecida a como la probabilidad antes descrita mide la verosimilitud que merecen los posibles sucesos inciertos.

La utilidad media de la decisión d_i, a la que llamaremos utilidad esperada de d_i y representaremos por $u^*(d_i)$, vendrá dada por

$$u^*(d_i) = \sum_{i=1}^{m_i} u(c_{ij}) p(\theta_{ij}|d_i, H)$$

Resulta natural que para elegir como decisión más razonable aquella que maximiza la itilidad esperada esto es aquella que hace máxima la expresión anterior entre las $k$alternativas$\{d_1, d_2, ..., d_k\}$. Este es el criterio de Bayes para la toma de decisiones.

2.3 Principios de coherencia

Formalmente, llamaremos una ópción y denotaremos por $l = \{c_1|A_1, c_2|A_2, ..., c_k|A_k\}$ a una situación en la que se obtiene la consecuencia c_1 si sucede A_1, la c_2 si sucede A_2, ..., la c_k si sucede A_k, con la codición de que los sucesos A_i sean exhaustivos y mutuamente excluyentes.

COMPARABILIDAD

Para todo par de opciones l_1 y l_2 es cierta una de las tres relaciones $l_1 < l_2$, $l_1 > l_2$ o $l_1 \sim l_2$. Además, es posible encontrar dos consecuencias c^* y c_* tales que $c^* > c_*$ y que para toda consecuencia c, $c_* \leq c \leq c^*$.

TRANSITIVIDAD

Si $l_1 > l_2$ y $l_2 > l_3$, entonces la $l_1 > l_3$. Analogamente si $l_1 \sim l_2$ y $l_2 \sim l_3$, entonces $l_1 \sim l_3$.

SUSTITUCIÓN Y DOMINANCIA

Si $l_1 > l_2$ cuando sucede A y $l_1 > l_2$ cuando sucede $\bar{A}$, entonces $l_1 > l_2$. Analogamente, si $l_1 \sim l_2$ cuando sucede A y $l_1 \sim l_2$ cuando sucede $\bar{A}$, entonces $l_1 \sim l_2$.

SUCESOS DE REFERENCIA

El decisor puede concebir un procedimiento de generar un punto aleatorio z en el cuadrado unidad, esto es, un número $z = (x, y)$, $0 \leq x \leq 1$, $0 \leq y \leq 1$ tal que para cualquier par de regiones R_1, R_2 del cuadrado unidad el suceso $\{z \in R_1\}$ le resulta menos verosímil que el suceso $\{z \in R_1\}$ si, y solo si, el área de R_1 es menor que la de R_2

2.4 Probabilidad como grado de creencia

Obviamente, la mayor parte de los sucesos inciertos que intervienen en un problema de decisión no presentan simetrías ni pueden ser repetidos indefinidamente en idénticas condiciones. Así por ejemplo, para decidir si debe ser utilizado un nuevo fármaco en un determinado paciente, hay que estimar la probabilidad de que este nuevo fármaco le sea eficaz y no le produzca transtornos secuendarios; sin embargo, no existen simetrías ni experiencia previa en idénticas condiciones que permitan específicar la probabilidad buscada.

La probabilidad de un suceso E en las condiciones H, que denotaremos por $p(E|H)$, es igual al área de la región R del cuadrado unidad elegida de forma que las opciones $l_1 = \{c^*|E, c_*\bar{E}\}$ y $l_2\{c^*|R, c_*|\bar{R}\}$ sen igualmente deseables en las condiciones H.

Cualquiera que sean las condiciones de referencia H la medida de probabilidad verifica:

i. Para todo suceso A, $0 \leq p(A|H) \leq 1$ y $p(H|H) = 1$

ii. Si A y B son dos sucesos incompatibles dado H,

$$p(A \cup B|H) = p(A|H) + p(B|H)$$

iii. Para todo suceso A y B,

$$p(A \cap B|H) = p(A|H) - p(B|A, H) = p(B|H) - p(A|B, H)$$

Para nosotros, la probabilidad $p(E|H)$ de un suceso E en las condiciones H es una medida, sobre la escala $[0, 1]$ de la verosimilitud que el decisor concede al suceso E en la situación descrita por H, esto es, una medida del grado de creencia en la ocurrencia de E que la información contenida en H le sugiere al decisor.

Hay dos puntos importantes que deben subrayarse. En primer lugar, la probabilidad asignada a un suceso es siempre condicional a la información que se posee sobre él; no existen probabilidades absolutas. En segunda lugar, estamos intentando cuantificar mediante probabilidades la información que una persona determinada posee sobre los sucesos inciertos que afectan a las consecuencias de sus decisiones: no existen probabilidades objetivas.

2.5 Maximzación de la utilidad esperada

La utilidad de una consecuencia c, que denotamos por $u(c)$ es la probabilidad x que debe asignarse a la mejor consecuencia c^* para que la consecuencia c se igualmente deseable que la opción $\{c^*|x, c_*|1-x\}$. De esta forma, para toda consecuencia c,

$$c \sim \{c^*|u(c), c_*|1 - u(c)\}$$

(CRITERIO DE DECISIÓN BAYES) Considérese el problema de decisión definido por $D = \{d_1, d_2, ..., d_k\}$ donde

$$d_i = \{c_{i1}|\theta_{i1}, c_{i2}|\theta_{i2}, ..., c_{im}|\theta_{im}\}$$

Sea $p(\theta_{ij}|d_i, H)$ la probabilidad de que suceda θ_{ij} si se elije d_1 en las condiciones H y sea $u(c_{ij})$ la utilidad de la consecuencia a que ello da lugar. Entonces, la utilidad esperada de la decisión d_i es

$$u^*(d_i) = \sum_{j=1}^{m_i} u(c_{ij}) p(\theta_{ij}|d_i, H)$$

y la decisión óptima es aquella con máxima utilidad esperada

2. Un médico cree que su paciente tiene una y sólo una de las enfermedades θ_1 , θ_2 , θ_3 , θ_4 y que sus probabilidades respectivas son $p(\theta_1) = 0.2$, $p(\theta_2) = 0.5$, $p(\theta_3) = 0.2$, $p(\theta_4) = 0.1$. El médico dispone de tres tratamientos alternativos t_1 , t_2 , t_3 cuya efectividad viene reflejada en la siguiente tabla:

tratamiento	theta1	theta2	theta3	theta4
t1	1	0	1	1
t2	0	1	0	1
t3	0	1	1	0

donde un uno en la fila i y columna j significa que el tratamiento t_i es efectivo contra la enfermedad θ_j , y un cero significa que no lo es. Igualando utilidad a efectividad, determinar el tratamiento óptimo.

```
p_theta1 <- 0.2
p_theta2 <- 0.5
p_theta3 <- 0.2
p_theta4 <- 0.1

e_t1 <- 1*p_theta1 + 0*p_theta2 + 1*p_theta3 + 1*p_theta4
e_t2 <- 0*p_theta1 + 1*p_theta2 + 0*p_theta3 + 1*p_theta4
e_t3 <- 0*p_theta1 + 1*p_theta2 + 1*p_theta3 + 0*p_theta4

e_t1
```

```
[1] 0.5
```

```
e_t2
```

```
[1] 0.6
```

```
e_t3
```

```
[1] 0.7
```

Por lo tanto el tratamiento t3 es el óptimo

Considérese que si un tratamiento no es efectivo puede procederse a otro y determínese la sucesión óptima de tratamientos, suponiendo que sus costos respectivos son $c(t_1) = 0.1$, $c(t_2) = 0.2$, $c(t_3) = 0.3$ en unidades de utilidad.

El conjunto de estados de la nauraleza es: $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$**, y asumimos que sólo uno de los estado puede ser cierto**

El conjunto de decisiones es: $D = \{t_1, t_2, t_3, t_{12}, t_{13}, t_{21}, t_{23}, t_{31}, t_{32}, t_{231}, t_{312}\}$**, esto es porque sólamente si el estado de la naturaleza es** θ_1**, sólo el tratamiento 1 será efectivo.**

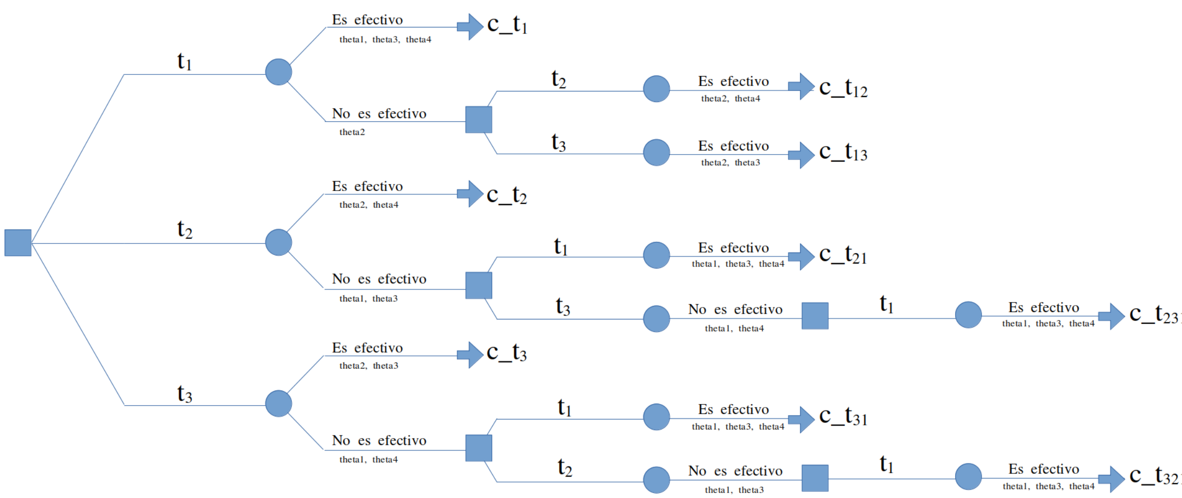

Figure 1: arbol ejercicio 2

Calculando la utilidad esperada de empezar por cada uno de los tratamientos:

```
c_t1 <- 0.1
c_t2 <- 0.2
c_t3 <- 0.3


c_t1 = p_theta1*c_t1 + p_theta3*c_t1 + p_theta4*c_t1
c_t2 = p_theta2*c_t2 + p_theta4*c_t2
c_t3 = p_theta2*c_t3 + p_theta3*c_t3
```

```r
c_t1
```

```
[1] 0.05
```

```r
c_t2
```

```
[1] 0.12
```

```r
c_t3
```

```
[1] 0.21
```

```r
c_t12 = c_t1 + p_theta2*c_t2
c_t13 = c_t1 + p_theta2*c_t3
c_t12
```

```
[1] 0.11
```

```r
c_t13
```

```
[1] 0.155
```

```r
c_t23 = c_t2 + p_theta1*c_t3 + p_theta3*c_t3
c_t21 = c_t2 + p_theta1*c_t1 + p_theta3*c_t1
c_t23
```

```
[1] 0.204
```

```r
c_t21
```

```
[1] 0.14
```

```
c_t31 = c_t3 + p_theta1*c_t1 + p_theta4*c_t1
c_t32 = c_t3 + p_theta1*c_t2 + p_theta4*c_t2
c_t31
```

```
[1] 0.225
```

```
c_t32
```

```
[1] 0.246
```

```
c_t231 = c_t23 + p_theta1*c_t1
c_t321 = c_t32 + p_theta1*c_t1
c_t231
```

```
[1] 0.214
```

```
c_t321
```

```
[1] 0.256
```

Lo más eficiente es empezar por el tratamiento 1, y si no funciona seguir con el 2

3. Sea $X \sim P(\theta)$, $\Theta = (0, \infty)$ y $A = [0, \infty)$. La función de pérdida es $L(\theta, a) = (\theta - a)^2$. Considerar reglas de decisión de la forma $\delta_c(x) = cx$. Suponer $\pi(\theta) = e^{-\theta}$ como la densidad inicial.

a. Calcular $\rho(\pi, a)$ y encontrar la acción de Bayes.

$$\rho(\pi, a) = \int_\Theta L(\theta, a) dF^{\pi^*}(\theta) = \int_0^\infty L(\theta, a)\pi(\theta)d\theta = \int_0^\infty (\theta - a)^2 e^{-\theta} d\theta = \int_0^\infty e^{-\theta}(\theta^2 + a^2 - 2a\theta)^2 d\theta$$

$$\rho(\pi, a) = \int_0^\infty e^{-\theta}\theta^2 d\theta + a^2 \int_0^\infty e^{-\theta} d\theta - 2a \int_0^\infty e^{-\theta}\theta^2 d\theta = 2 + a^2 - 2a$$

b. Encontrar $R(\theta, \delta_c)$.

$$R(\theta,\delta_c) = E_\theta^X[L(\theta,\delta(X))] = E_\theta^X[L(\theta,\delta_c(X))] = E_\theta^X[L(\theta,cX)] = E_\theta^X[(\theta-cX)^2]$$
$$R(\theta,\delta_c) = E_\theta^X[(\theta-cX+c\theta-c\theta)^2]$$
$$R(\theta,\delta_c) = E_\theta^X(c[\theta-X]+[1-c]\theta)^2$$
$$R(\theta,\delta_c) = E_\theta^X[c^2(\theta-X)^2] + 2E_\theta^X[c(\theta-X)(1-c)\theta] + E_\theta^X[(1-c)^2\theta^2]$$
$$R(\theta,\delta_c) = c^2E_\theta^X[(\theta-X)^2] + 2c(1-c)\theta E_\theta^X[(\theta-X)] + (1-c)^2\theta^2E_\theta^X[1]$$

Sabemos que $E_\theta^X[(\theta-X)] = \theta E_\theta^X[1] - E_\theta^X[X] = \theta-\theta = 0$ Y también que la varianza es $E_\theta^X[(\theta-X)^2] = \theta$, por lo que

$$R(\theta,\delta_c) = c^2\theta + (1-c)^2\theta^2$$

c. Mostrar que δ_c es inadmisible si $c > 1$

Si $c > 1$, entonces $c^2\theta + (1-c)^2\theta^2 > 1$ por lo que basta encontrar $R(\theta,\delta_1)$ tal que para todo $\theta \in \Theta$,$R(\theta,\delta_1) \leq R(\theta,\delta_c)$. Si $R(\theta,\delta_1) = 1$ esta condición se cumple

d. Encontrar $r(\pi,\delta_c)$.

$$r(\pi,\delta_c) = E^\pi[R(\theta,\delta_c)] = E^\pi[c^2\theta+(1-c)^2\theta^2] = c^2E^\pi[\theta]+(1-c)^2E^\pi[\theta^2] = c^2\theta+(1-c)^2\theta$$

e. Encontrar el valor de c que minimiza $r(\pi,\delta_c)$.

$$\frac{d}{dc}(c^2\theta + (1-c)^2\theta) = 2c\theta - 2(1-c)\theta = 2c\theta - 2\theta + 2c\theta = 4\theta c - 2\theta = 0$$

$$c = 1/2$$

4. La Junta de Gobierno del Banco de México se enfrenta con una de las siguientes tres acciones:

- a1 : incrementar la tasa de interés 25 puntos base (pb);
- a2 : mantener la tasa de interés en el nivel actual;
- a3 : disminuir la tasa de interés en 25pb.

Dependiendo de los resultados de la medición de inflación si sube (θ_1) o se mantiene (θ_2) o baja (θ_3), se espera incurrir en los siguientes costos monetarios:

caso	a1	a2	a3
theta1	-10	-5	-3
theta2	-5	-5	-2
theta3	1	0	-1

a. Determinar si cada acción es admisible o inadmisible.

Toda las acciones son admisibles, dado que dependiendo el estado de la naturaleza hay una acción que es igual o más de conveniente que las otras.

b. La JG cree que θ tiene distribución de probabilidad $\pi(\theta) = 0.2I(\theta = \theta_1) + 0.3I(\theta = \theta_2) + 0.5I(\theta = \theta_3)$. Ordenar las acciones de acuerdo a su pérdida esperada Bayesiana y encontrar la decisión de Bayes.

```
theta1 <- 0.2
theta2 <- 0.3
theta3 <- 0.5

c_a1 <- -10*theta1 -5*theta2 +1*theta3
c_a2 <- -5*theta1 -5*theta2 +0*theta3
c_a3 <- -3*theta1 -2*theta2 -1*theta3

c_a1
```

```
[1] -3
```

```
c_a2
```

```
[1] -2.5
```

```
c_a3
```

```
[1] -1.7
```

Se toma la acción a_1