Estadística

Apuntes de repaso

Daniel Pérez Efremova

Curso 2021-2022

Índice

1. Datos

1.1. Descripción de una sola variable

1.1.1. Frecuencias

La frecuencia absoluta n_i de un suceso x_i es el número de veces que éste se observa en la muestra. Por otro lado, si la muestra tiene n observaciones, la frecuencia relativa de x_i es el número de veces que se ha observado x_i con respecto al total de observaciones en la muestra: $f_i = n_i/n$.

Ejemplo 1 *Poner ejemplo aquí*

1.2. Descripción de varias variables

2. Modelos

2.1. Probabilidad y variables aleatorias

El objetivo de la probabilidad es inferir las propiedades de una población a partir de una muestra. El intrumento que permite dichas inferencias es un **modelo de la población**. El cálculo de probabilidades permite medir la incertidumbre de un suceso según el modelo de la población.

2.1.1. Probabilidad y sus propiedades

En esta sección se explican los conceptos más relevantes de la probabilidad y que son de uso común en toda la asignatura.

Población, muestra y experimento

Estos conceptos son la base para referirse a los distintos elementos de un estudio estadístico y no deben confundirse.

Definición 1 *(población) Una población es un conjunto de elementos de los que se quiere extraer o inferir cierta información. Si la población contiene un número finito de elementos se dice que la población es finita. Por otro lado, si la cantidad de elementos no se puede determinar de manera exacta se asume que es infinita.*

Definición 2 *(experimento) Un experimento es la acción de observar una determinada característica en un elemento de la población en estudio.*

Definición 3 *(muestra) Una muestra es un conjunto de experimentos llevados a cabo en la población en estudio.*

Estos conceptos se entienden mejor con un ejemplo sencillo.

Ejemplo 2 *Supongamos que se quiere averiguar qué cantidad de hombres y mujeres hay en la ciudad de Madrid. En este caso, la población en estudio son todas las personas de Madrid. Un experimento es observar a una persona concreta y comprobar si es hombre o mujer. Una muestra son varios experimentos donde se observan distintas personas para comprobar si son hombre o mujer.*

Son importantes porque al estudiar los modelos de probabilidad, se tiene una población modelo sobre la que se hacen inferencias a partir de muestras. Además, un modelo define su ley de probabilidad a través de experimentos.

Suceso y Espacio muestral

Estos conceptos son importantes para definir claramente cuáles son todos los posibles resultados de un experimento y qué probabilidad se está calculando.

Definición 4 *(Suceso) Un suceso es el resultado de un experimento. Además, si un suceso se expresa como unión o intersección de varios sucesos se dice que es un suceso compuesto. Los sucesos se suelen denotar con una letra en mayúscula.*

Definición 5 *(espacio muestral) El espacio muestral es el conjunto de todos los sucesos posibles. Siempre tiene que ocurrir alguno de ellos y son mutuamente excluyentes. Suele denotarse con* Ω.

Definición 6 *(Suceso complementario) Dado un suceso* $A \in \Omega$, *se dice que el suceso* $\overline{A} = \Omega - A$ *es el suceso complementario de* A.

Así, cuando se hace un estudio estadístico, se define una población y un espacio muestral que se espera encontrar en ella. Además, cuando se calculan probabilidades, se tiene que decir explícitamente cuál es el suceso al que se le calcula la probabilidad.

Ejemplo 3 *Supongamos que tenemos un dado equilibrado de 6 caras. Al lanzar el dado, los posibles resultados son* $\{1, 2, 3, 4, 5, 6\}$. *El espacio muestral es* $\Omega = \{1, 2, 3, 4, 5, 6\}$ *y esta compuesto por los sucesos elementales* $A_i =$ *salir el número* i. *Un suceso compuesto podría ser* $A = A_1 \cup A_2 = \{1, 2\}$. *Por otro lado, el suceso complementario de* $A_1 = \{1\}$ *es* $\overline{A_1} = \Omega - \{1\} = \{2, 3, 4, 5, 6\}$

Probabilidad[1]

La probabilidad mide la incertidumbre asociada a un suceso concreto respecto a todo el espacio muestral. En esencia, no es más que una función que asigna a cada suceso un número, y este número indica cuánto de común es observar el suceso.

[1] En las diapositvas esta parte se explica usando un enfoque frecuentista, es decir, usando las frecuencias n_i. En estos apuntes se sigue un enfoque axiomático (Kolmogorov).

Definición 7 *(Probabilidad) Una probabilidad P es una función que asigna a cada suceso del espacio muestral Ω un número real entre 0 y 1.*[2]

A continuación, se define la regla más comun e intuitiva para calcular probabilidades bajo el supuesto de que todos los sucesos elementales del espacio muestral son equiprobables, es decir, que realizado un experimento, la probabilidad de observar cualquier suceso elemental es la misma.

Definición 8 *(Regla de Laplace) Dado un espacio muestral Ω y un suceso A, la probabilidad del suceso viene dada por:*

$$P(A) = \frac{n^{o} \text{ de sucesos favorables a } A}{\text{Número total de sucesos posibles}}$$

Ejemplo 4 *Continuando con el ejemplo 3, si el dado esta equilibrado, la probabilidad de cualquier suceso elemental es*

$$P(A_i) = \frac{n^{o} \text{ de resultados en los que sale el número } i}{n^{o} \text{ de resultados posibles}} = \frac{1}{6}$$

Por otro lado, la probabilidad del suceso compuesto $A_1 \cup A_2$ es

$$P(A_1 \cup A_2) = \frac{n^{o} \text{ de resultados en los que sale el número } 1 \text{ o } 2}{n^{o} \text{ de resultados posibles}} = \frac{2}{6} = \frac{1}{3}$$

Además, es importante tener claras las propiedades de las probabilidades para no caer en absurdos y detectar errores de cálculo. Éstas se definen a partir de los axiomas de Kolmogorov, que son un conjunto de propiedades que deben cumplirse siempre que se calculen probabilidades.

Definición 9 *(Axiomas de Kolmogorov) Dado un espacio muestral Ω y una función de probabilidad P, se cumple que:*

1. *$\forall A \in \Omega$ se tiene que $P(A) \in [0,1]_{\mathbb{R}}$*
2. *$P(\Omega) = 1$*
3. *Si $\{A_1, \ldots, A_n\}$ son sucesos elementales, entonces $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$*

El primer axioma dice que no pueden haber probabilidades negativas ni mayores que 1. El segundo dice que en cualquier experimento siempre debe ocurrir algun suceso del espacio muestral. El tercero dice que dado un conjunto de sucesos elementales, la probabilidad de que ocurra alguno de ellos es la suma de sus probabilidades[3].

De este conjunto de axiomas se deducen todas las propiedades de las probabilidades (las demostraciones se dejan en el anexo):

[2] En el entorno técnico es una aplicación $P : \Omega \longrightarrow [0,1]_{\mathbb{R}}$

[3] Nótese que para que se cumpla los sucesos deben ser elementales, esto es, se verifica que $A_i \cap A_j = \emptyset$ para cualesquiera i, j.

Definición 10 *(propiedades de las probabilidades) Las probabilidades cumplen las siguientes propiedades:*

1. *(Probabilidad del suceso vacío)* $P(\emptyset) = 0$

2. *(Monotonía) Dados dos sucesos* $A, B \in \Omega$*, si* $A \subseteq B \Rightarrow P(A) \leq P(B)$

3. *(Sucesos definidos segun otro suceso) Dados dos sucesos* $A, B \in \Omega$ *se tiene que* $P(A) = P(A - B) + P(A \cap B)$

4. *(Regla del complementario) Para cualquier suceso* $A \in \Omega$*, se verifica que* $P(\overline{A}) = 1 - P(A)$.

5. *(Probabilidad de la unión) Dados dos sucesos* $A, B \in \Omega$ *se tiene que* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

6. *(Leyes de Morgan) Dados dos sucesos* $A, B \in \Omega$ *se tiene que*

$$P(\overline{A \cup B}) = P(\overline{A} \cap \overline{B})$$

$$P(\overline{A \cap B}) = P(\overline{A} \cup \overline{B})$$

La primera propiedad dice que la probabilidad de un suceso vacío es 0, lo que concuerda con que siempre debe ocurrir algún suceso del espacio muestral en los experimentos. La segunda dice que si un suceso es parte de otro suceso más amplio, entonces la probabilidad del suceso más amplio acota al resto[4]. La tercera dice que un suceso A (y por tanto, su probabilidad) puede definirse en base a otro suceso B como: *lo que hay en* A *y además está* B *más lo que hay en* A *y no está en* B. La regla del complementario no es más que una aplicación de la tercera al caso en que $B = \Omega$. La quinta propiedad evita contar dos veces la intersección de los dos sucesos A, B. Las leyes de Morgan permiten calcular la probabilidad de sucesos complementarios de una manera rápida y tienen su origen en la lógica proposicional.

2.1.2. Probabilidad condicionada

En esta sección se los resultados más notables de probabilidades condicionadas, se introduce el concepto de dependencia de sucesos y se enuncia el teorema de Bayes con algunos ejemplos.

Definición 11 *(Probabilidad condicionada) Dado un espacio muestral* Ω *y una función de probabilidad* P*, entonces, la probabilidad de un suceso* $A \subset \Omega$ *condicionada por otro suceso* $B \subset \Omega$ *se define como*

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

[4]Nótese que cualquier suceso verifica $A \subset \Omega$, luego $P(A) \leq P(\Omega) = 1$ y se obtiene que la probabilidad está acotada superiormente a 1. Así que esta propiedad no es más que una extensión a cualquier par de subconjuntos A, B de la acotación en 1 de las probabilidades.

Esta probabilidad mide la incertidumbre asociada a un suceso A cuando también ocurre otro suceso B. Ligada a este concepto está la noción de independencia de sucesos.

Definición 12 *(Independencia de sucesos) Dado un espacio muestral Ω y una función de probabilidad P, los sucesos $A, B \subset \Omega$ se dicen independientes si se cumple que*

$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$

Estas dos condiciones pueden resumirse en una única expresión más práctica

$$P(A \cap B) = P(A)P(B)$$

Es decir, dos sucesos son independientes si la ocurrencia de un suceso no altera la probabilidad de aparición del otro.

Teorema de Bayes

Ligado al concepto de probabilidad condicionada se encuentra el teorema de Bayes. La terminología usual de probabilidades a priori y posteriori nace de su aplicación en campos como la economía, medicina, etc. Primero se ve como llegar matemáticamente al resultado del teorema y después con un ejemplo se explica esta terminología usual.

Teorema 1 *Sea Ω un espacio muestral y P una función de probabilidad. Consideremos un conjunto de sucesos que cubre todo el espacio muestral i.e. $A = \{A_1, \ldots, A_n\}$ tal que $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$ y un suceso $B \subset \Omega$. Entonces, la probabilidad de un suceso $A_i \in A$ condicionada por el suceso B es:*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

Demostración

Como se cumple que $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$ se puede expresar B en relación al especio muestral Ω como sigue

$$B = B \cap \Omega = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$$

Además, como los sucesos elementales son independientes, aplicando probabilidades a ambos lados de la igualdad anterior se tiene que

$$P(B) = P((A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)) = \sum_{i=1}^{n} P(A_i \cap B) \qquad (1)$$

Ahora bien, se sabe por la definición de probabilidad condicionada (11) que

$$P(A_i \cap B) = P(A_i)P(B|A_i) \tag{2}$$

Del mismo modo se tiene que

$$P(B) = \frac{P(A_i \cap B)}{P(A_i|B)} \tag{3}$$

Aplicando 3 en el lado izquierdo de 1 y 2 sobre el lado derecho de la ecuación 1 se tiene:

$$\frac{P(A_i \cap B)}{P(A_i|B)} = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

Ordenando términos se tiene:

$$P(A_i|B) = \frac{P(A_i \cap B)}{\sum_{i=1}^{n} P(B)P(A_i|B)}$$

Por último, volviendo a aplicar 2 sobre la expresión anterior, se obtiene finalmente el resultado del teorema de Bayes:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

■

En la demostración del teorema se ve que el resultado nace de expresar la probabilidad de un suceso B a través de las condicionadas con un conjunto de sucesos $\{A_1, \ldots, A_n\}$.

Ahora bien, por lo general, este teorema se utiliza para resolver problemas de experimentos en dos fases.

Por ejemplo, supongamos que realizamos un experimento donde medimos si unos pacientes tienen un virus o no y si dan positivo o negativo en un test. Si primero se mide si tienen el virus y después se mide el resultado del test, el teorema de Bayes respondería a una pregunta del tipo: ¿Qué probabilidad tiene un paciente de tener el virus si se sabe que en el test ha dado negativo? Es decir, el teorema trata de responder a preguntas acerca de la primera fase cuando solo se conoce información de los resultados de la segunda fase.

Ejemplo 5 *(Aplicación del Teorema de Bayes) Siguiendo con el ejemplo anterior y con la misma notación que en el Teorema 1, los sucesos son:* $A_1 =$ *"tener el virus" con* $P(A_1) = 0{,}1$*;* $A_2 =$ *"no tener el virus" con* $P(A_2) = 0{,}9$*;* $B =$ *"test negativo" con probabilidad desconocida. Además se sabe* $P(B|A_1) = 0{,}8$ *y* $P(B|A_2) = 0{,}2$*. Entonces, si sabemos que el test ha dado negativo, la probabilidad de que el paciente tenga el virus es:*

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)} = \frac{0{,}1 \times 0{,}8}{0{,}1 \times 0{,}8 + 0{,}9 \times 0{,}2} = 0{,}307$$

2.1.3. Variables aleatorias

2.2. Modelos de distribución de probabilidad

2.2.1. El proceso de Bernoulli

2.2.2. El proceso de Poisson

2.2.3. Las distribuciones de duraciones de vida

2.2.4. La distribución Normal

2.2.5. La distribución Log-Normal

2.3. Modelos Multivariantes

Anexo

Esta sección sirve de complemento para las secciones anteriores, profundizando en los conceptos introducidos.

Demostración de las propiedades de las probabilidades

Demostración del Teorema de Bayes