Probabilidad discreta

Motivo En el día anterior, se discutió la definición de un espacio de probabilidad. Una definición consecutiva es la noción de variable aleatoria. Se introduce como una función de las realizaciones elemental ω. Por ejemplo, consideramos un juego donde se lanzan dos dados. El resultado es una realización $\omega = (\omega_1, \omega_2) \in \{1, 2, 3, 4, 5, 6\}^2$. La suma S de los valores de los dos dados es un número que depende de ω. Hemos definido una función S que asocia a ω, el valor $\omega_1 + \omega_2$. Hablamos de variable aleatoria para designar S.

1 Partes discretas de $\mathbb{R}$ y definiciones de variables discretas

La probabilidad discreta se interesa al caso particular de un espacio muestral Ω discreto. En ese contexto, la definición de variable aleatoria es muy natural.

Definición 1. *Una parte $D \subset \mathbb{R}$ es **discreta** si cada punto $\omega \in D$ es aislado*[1].

Ejemplos

- Un conjunto finito de $\mathbb{R}$ es discreto. $(\star)$
- Los naturales $\mathbb{N}$ es un conjunto discreto.
- Los racionales en $[0, 1]$ no es un conjunto discreto. $(\star)$

Una parte discreta es en particular numerable[2] pero el contrario no es verdad. El concepto de espacio discreto da nacimiento naturalmente a la noción de variable aleatoria discreta.

Definición 2. *Por un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$, una función $X : \Omega\mathbb{R}$ es una **variable aleatoria discreta** si*

1. *$X(\Omega) = \{X(\omega) : \ \omega \in \Omega\}$ es discreto.*
2. *Para todo $x \in X(\Omega)$, $A(x) = \{\omega : \ X(\omega) = x\}$ es un evento de $\mathcal{A}$.*

Un abuso de notaciones común es de escribir $\underline{\mathbb{P}(X = x) \text{ por la probabilidad } \mathbb{P}(A(x))}$.

En consecuencia, los eventos $A(x)$ forman una partición[3] de Ω. Ademas, la probabilidad de la unión numerable de eventos disjuntos se convierte en una suma,

$$\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = \sum_{x \in X(\Omega)} \mathbb{P}(A(x)) = \mathbb{P}\left(\bigcup_{x \in X(\Omega)} A(x)\right) = 1.$$

En consecuencia, por una variable discreta X, los eventos $A(x) = \{X = x\}$ se llaman **eventos elementales** de la variable aleatoria X.

Ley de una variable aleatoria discreta La variable aleatoria X permite de transferir la probabilidad $\mathbb{P}$ de Ω sobre $\mathbb{R}$: se consideran los $\mathbb{P}(X = x)$ como masas puntuales ubicadas a los puntos x. La probabilidad de un conjunto cualquiera A de $\mathbb{R}$ es definido como la suma de las masas de los puntos adentro de A. Concretamente, denotamos $p_x = \mathbb{P}(X = x)$ y para $B \in \mathcal{P}(\mathbb{R})$[4], definimos

$$\mathbb{P}_X(B) = \sum_{x \in X(\Omega) \cap B} \mathbb{P}(X = x) = \sum_{x \in X(\Omega) \cap B} p_x.$$

La función $\mathbb{P}_X : \mathcal{P}(\mathbb{R}) \to [0, 1]$ se llama **ley** de la variable X.

[1] Un punto ω es aislado en D si existe intervalo centrado en ω que no contiene un otro punto de D que ω.
[2] Una parte D numerable de $\mathbb{R}$ es un conjunto que se puede enumerar $D = \{x_1, x_2, \dots\}$
[3] Son obviamente disjuntos y por definición la unión vale Ω
[4] $\mathcal{P}(\mathbb{R})$ es el conjunto de todos los subconjuntos de $\mathbb{R}$.

2 Distribuciones discretas clásicas

Definición de variables por su ley La noción de ley de una variable aleatoria X caracteriza la parte importante de la variable X. En efecto, en los experimentos, no nos importa conocer ω. Por ejemplo, el conocimiento de la suma de dos dados S no requiere saber exactamente el valor del dado 1 y del dado 2. Finalmente, es suficiente de definir los p_x sobre un conjunto discreto para definir la ley de una variable aleatoria. Es lo que haremos el lo siguiente. Cuidado, dos variables pueden tener la misma ley sin ser iguales. El ejemplo dela lanzamiento de dos dados permite ver que X el resultado del primer dado y Y el resultado del segunda dado tienen la misma ley pero no son iguales : $X(\omega) = Y(\omega)$ para todo ω. En efecto, decir que son iguales significa que solo se observaría dobles...

Ley de Bernoulli $\mathcal{B}(p)$ Una variable X de **Bernoulli de parametro** $p \in [0,1]$ tiene un soporte $\{0,1\}$ y es tal que $\mathbb{P}(X=1) = p$ y $\mathbb{P}(X=0) = 1-p$. Eso modela el resultado aleatorio de un experimento binario. Es muy usado para modelar una situación de éxito/fracaso.

Ley uniforme $U(F)$ Una variable X es **uniforme sobre el conjunto finito** F si $\forall x \in F, \mathbb{P}(X=x) = \frac{1}{|F|}$. Eso modela una situación donde no hay ninguna información a priori sobre el resultado de un experimento.

Ley binomiale $\mathcal{B}(n,p)$ Una variable X es **binomiale de parametros** $n \in \mathbb{N}$ y $p \in [0,1]$ si tiene un soporte $\{0,1\ldots,n\}$ y $\mathbb{P}(X=i) = \binom{n}{i}p^i(1-p)^{n-i}$. El parámetro p se llama probabilidad de éxito y n el número de intentos. Esa variable modela el número de éxito en una repetición de n experimentos idénticos y independientes.

Ley geométrica $\mathcal{G}(p)$ Una variable X es **geométrica de parámetro** $p \in [0,1]$ si tiene un soporte en $\mathbb{N}$ y $\forall i \in \mathbb{N}, \mathbb{P}(X=i) = (1-p)^{i-1}p$. El parámetro p se llama probabilidad de éxito. Esa variable modela el tiempo de primer éxito de un experimento que ocurre con probabilidad p.

Ley de Poisson $\mathcal{P}(\lambda)$ Una variable X es de **Poisson de parámetro** $\lambda > 0$ si tiene soporte en $\mathbb{N}\cup\{0\}$ y $\forall i \in \{0,1,\ldots\}, \mathbb{P}(X=i) = \frac{\lambda^i}{i!}e^{-\lambda}$. Esa variable sirve para modelar el número de eventos independientes que ocurren durante un tiempo dado.

3 Contextos y ejemplos de uso de las distribuciones clásicas

Justificación de la ley binomiale Supongamos que se denotan $\varepsilon_1,\ldots,\varepsilon_n$, n variables aleatorias de Bernoulli de parámetro p. El número de éxitos N se relaciona con las variables anteriores gracias a la relación

$$N = \sum_{i=1}^{n} \varepsilon_i.$$

Los valores posible de N son $\{0,\ldots,n\}$. Calculamos las probabilidades $p_k = \mathbb{P}(N=k)$. El vector $(\varepsilon_1,\ldots,\varepsilon_n)$ puede tomar todos los valores en $\{0,1\}^n$. Los vectores que contienen k unos y $n-k$ zeros son equiprobables[5] de probabilidad $p^k(1-p)^{n-k}$. El evento $\{N=k\}$ es la unión disjunta de los eventos elementales que acabamos de describir. Pero el número de vectores teniendo exactamente k unos corresponde al número de parte de cardinal k en un conjunto de n elementos. Vimos que vale $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. A fin de cuenta,

$$\mathbb{P}(N=k) = \binom{n}{k} p^k(1-p)^{n-k}.$$

Justificación de la ley geométrica Para calcular la probabilidad $\mathbb{P}(X=k)$, una posibilidad es de escribir el evento $\{X=k\}$ con eventos mas sencillos. Sean $R_i = \{$éxito al i-ésimo intento$\}$. Podemos escribir

$$\{X=k\} = \left(\bigcap_{i=1}^{k-1} R_i^c\right) \cap R_k.$$

[5] Tienen la misma probabilidad de aparición

Tenemos inmediatamente $\mathbb{P}(R_i) = p$ y entonces $\mathbb{P}(X = k) = (1-p)^{k-1}p$. Podemos verificar que $\sum_{i=1}^{\infty} \mathbb{P}(X = i) = 1$.
Se puede probar de la misma manera que $\mathbb{P}(X > n) = (1-p)^n$. $(\star)$

Propiedades de la ley de Poisson La descomposición analítica de la función exponencial se escribe

$$e^{\lambda} = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!},$$

lo que nos asegura que $\sum_{i=0}^{\infty} \mathbb{P}(X = i) = 1$. Una propiedad famosa de las variables de Poisson es que aproximan bien las variables binomiales cuando n es grande y p es pequeño. Mas concretamente, consideramos una sucesión p_n de parámetros tal que la sucesión np_n tiene un limite $\lambda > 0$ (cuando $n \to +\infty$). Denotamos X_n una variable binomiale $\mathcal{B}(n, p_n)$ y Y una variable de Poisson de parámetro λ. Entonces

$$\mathbb{P}(X_n = k) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(Y = k).$$

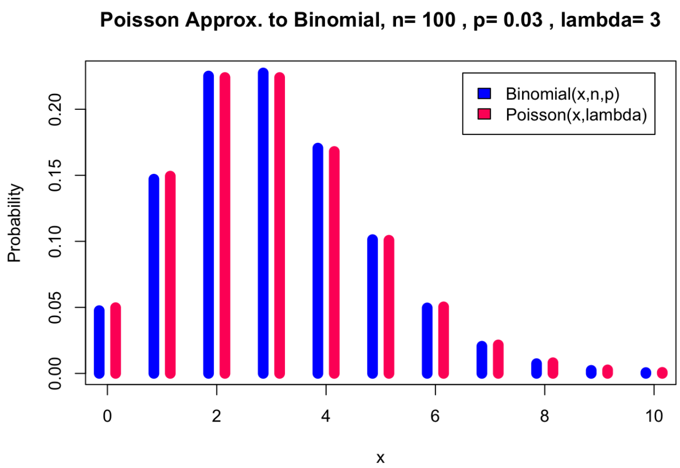
En efecto,

$$\mathbb{P}(X_n = k) = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} \times \frac{n!}{n^k (n-k)!} \times \left(\frac{np_n}{\lambda}\right)^k \times (1-p_n)^{n-k}.$$

Como

$$\frac{n!}{n^k(n-k)!} \to 1, \qquad \left(\frac{np_n}{\lambda}\right)^k \to 1, \qquad (1-p_n)^{n-k} \to e^{-\lambda}$$

obtenemos el resultado.



4 Función de distribución y sus propiedades básicas

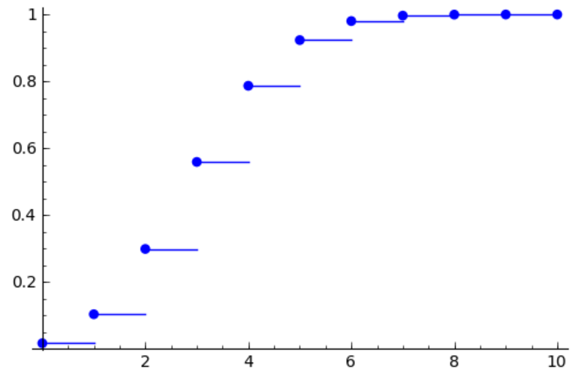
La función de distribución es una representación única de la ley de una variable aleatoria. Es una herramienta visual y de cálculo poderoso.

Definición 3. *La **función de distribución** de una variable aleatoria discreta X es la función $F_X : \mathbb{R} \to [0, 1]$ y que a cada real t asocia la probabilidad $\mathbb{P}(X \leq t)$.*

Algunas propiedades de F_X Siempre se cumplen las propiedades

- Le función F_X es no-decreciente. $(\star)$
- F_X es continua a la derecha y tiene un límite a la izquierda.$(\star)$
- $\lim_{t\to-\infty} F_X(t) = 0$ y $\lim_{t\to+\infty} F_X(t) = 1$ $(\star)$

Para la variables discretas la función de distribución es constante por pedazos.



Teorema 1. *Si dos variables aleatorias discretas X y Y tienen la misma función de distribución, entonces X y Y tienen la misma ley de probabilidad.*

Dem. Primero, probamos que para cada $x \in \mathbb{R}$,

$$\{X = x\} = \bigcap_{n \in \mathbb{N}} \left\{ x - \frac{1}{n} < X \leq x \right\}.$$

Claramente, el primer evento está incluido en el segundo. Para ver el sentido contrario, tomando ω en la intersección de la derecha, se observa que $\forall n \in \mathbb{N}$, $x - \frac{1}{n} < X(\omega) \leq x$. El $X(w)$ (que no depende de n) está encerrado entre dos sucesiones de límite x cuando $n \to \infty$, por lo tanto $x \leq X(\omega) \leq x$ y $X(\omega) = x$. La vez pasada, vimos que si B_n es una sucesión decreciente de eventos tal que $B_n \downarrow B$, entonces $\mathbb{P}(B_n) \to \mathbb{P}(B)$. Entonces, considerando $B_n = \left\{ x - \frac{1}{n} < X \leq x \right\} = \bigcap_{k=1}^{n} \left\{ x - \frac{1}{k} < X \leq x \right\}$, tenemos $B_n \downarrow \{X = x\}$. Por lo tanto

$$\begin{aligned} \mathbb{P}(X = x) = \lim_{n \to \infty} \mathbb{P}(B_n) &= \lim_{n \to \infty} F_X \left(x - \frac{1}{n} \right) - F_X(x) \\ &= \lim_{n \to \infty} F_Y \left(x - \frac{1}{n} \right) - F_Y(x) = \mathbb{P}(Y = x). \end{aligned}$$

Como eso es verdad para todo $x \in \mathbb{R}$, lo es para todo x en el soporte de X y entonces, las dos variables X y Y tienen la misma ley de probabilidad. $\square$

5 Problemas

1. Sea X una variable aleatoria tal que $P(X = 2) = P(X = 3) = 1/10$ y $P(X = 5) = 8/10$.

   1.1. Grafique la función de distribución $F(x)$.

   1.2. Use F para encontrar $P(2 < X \leq 4.8)$ y $P(2 \leq X \leq 4.8)$

2. ($\star$)Suponga que en un lago hay 1000 peces, suponga además que 50 de ellos fueron atrapados, marcados y regresados al lago y que ha pasado un tiempo suficiente para que esos 50 se mezclen aleatoriamente en la población.

   2.1. Obtenga una expresión para $P(x)$, la probabilidad de obtener x peces marcados, en una muestra aleatoria de 10 peces.

   2.2. Calcule $P(0), P(1)$ y $P(2)$ y compare estos valores con los que se obtienen usando una distribución Binomial.

3. Suponga que las hembras de cierta especie de insectos ponen X huevos durante un periodo de tiempo (digamos 1 mes). Suponga además que es razonable modelar el comportamiento de la variable aleatoria X mediante un distribución Poisson con media λ. Ahora, cada huevecillo tiene una probabilidad p de exitosamente desarrollar un nuevo individuo.

3.1. Encuentre la distribución del número de huevos, Y, que se desarrolla exitosamente durante un periodo de tiempo para una hembra dada.

Ayuda:

$$\mathrm{P}(Y = y) = \sum_{x=y}^{\infty} \mathrm{P}(Y = y|X = x)\,\mathrm{P}(X = x)$$

3.2. La distribución anterior es una distribución Poisson, ¿cuál es su media? (es interesante - y lógico- que si el número de huevos es Poisson, entonces el número de nuevos individuos también es Poisson)

4. ($\star$)Supongamos que una moneda, al ser lanzada, tiene probabilidad p de caer cara. La moneda es lanzada hasta la primera vez que aparece sol. Sea X el número total de caras, ¿ Qué es $\mathbb{P}(X > m)$? Encontrar la función de distribución de la variable aleatoria X.

5. Las aerolíneas han calculado la probabilidad de que cada individuo que reserva un asiento y no se presenta es de 1/10 independientemente de la decisión de algún otro pasajero. Se sabe que Volare siempre vende 10 boletos para los 9 asientos de sus aviones mientras Internacional siempre vende 20 boletos para los 18 asientos de sus aviones. ¿ Cuál de ambas compañias se sobre vende mas seguido?

6. Muestre que si X es una variable aleatoria binomial o Poisson, entonces sus función de masas de probabilidades $f(k) := \mathbb{P}(X = k)$ satisface $f(k-1)f(k+1) \leq f(k)^2$.