Hoy demostrarmos que: La mala suerte no es verdad

Bioestadística (Clase 5):

Variables aleatorias discretas

Ernesto Mordecki

CMAT, Facultad de Ciencias, Universidad de la República.

Uruguay

Contenidos de la clase

- Independencia de sucesos
- Variables aleatorias discretas
  - Variable aleatoria uniforme
  - Variable aleatoria de Bernoulli
  - Variable aleatoria Binomial
- Paseo al azar, triangulo de Pascal y variable aleatoria Binomial
- Variable geométrica

Sucesos independientes

- Decimos que dos sucesos A y B son independientes cuando se verifica

$$\boxed{P(A \cap B) = P(A) \times P(B).}$$

- La independencia es un concepto matemático.

Ejemplo

Supongamos que tiramos dos dados (el negro y el rojo)

- Veamos que el resultado del dado negro es independiente del dado rojo si N, R son los resultados respectivos.
- El primer suceso es $N = k$ (el dado negro toma el valor k)
- El segundo suceso es $R = j$ (el dado rojo toma el valor j)
- Tenemos

$$P(N = k, R = j) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = P(N = k)P(R = j)$$

- Decimos que los resultados N y R son independientes.

Ejemplo de sucesos dependientes

Consideremos los sucesos:

- $A = \{\text{la suma es mayor estricta que } 7\} = \{N + R > 7\}$
- $B = \{\text{el dado negro es mayor o igual a } 4\} = \{N \geq 4\}$
- Veamos si estos sucesos A y B son independientes

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Cuadro: En azul la suma mayor que 7, a la derecha de la línea el dado negro mayor o igual a 4

$$P(A) = \frac{15}{36}, \quad P(B) = \frac{18}{36} = \frac{1}{2}, \quad P(A \cap B) = \frac{12}{36} = \frac{1}{3}$$

$$P(A)P(B) = \frac{15}{72} \neq P(A \cap B) = \frac{1}{3} = \frac{24}{72}.$$

Los sucesos *A* y *B* no son idependientes.

Variables aleatorias discretas

- El ejemplo anterior nos sirvió para introducir las variables aleatorias R y N
- Una variable aleatoria es una función que nos lleva del espacio Ω a $\mathbb{R}$
- Como estamos en $\mathbb{R}$ nos permite utilizar herramientas de cálculo
- La variable aleatoria es discreta cuanto toma un conjunto finito $\{0, 1, 2, \ldots, N\}$ o numerable $\{0, 1, 2, \ldots\}$ de valores.
- Nos da flexibilidad para considerar modelos de probabilidad no equiprobables.

Variable aleatoria uniforme

- El primer ejemplo que consideramos es una variable aleatoria X que toma valores $1, \ldots N$.
- Asumimos que todos los valores son equiprobables.
- Tenemos

$$P(X = k) = \frac{1}{N}, \quad \text{para } k = 1, \ldots, N$$

- Este modelo:
  - Si $N = 2$ corresponde a tirar una moneda equilibrada, y asignar 1 y 2 a cara y número.
  - Si $N = 6$ corresponde a tirar un dado equilibrado.

Variable aleatoria de Bernoulli

- Nuestra variable toma dos valores:

$$X = \begin{cases} 0, & \text{que corresponde al fracaso} \\ 1, & \text{que corresponde al éxito} \end{cases}$$

  de un experimento aleatorio con dos resultados posibles

- Los valores **no** son equiprobables. Para un $p \in (0,1)$:

$$P(X=0) = 1-p, \qquad P(X=1) = p.$$

- Este modelo se puede aplicar en muchas situaciones:
  - Si un producto es defectuoso o no,
  - Si salvo un examen o lo pierdo,
  - Si en un texto en un lugar tengo una vocal o una consonante . . . etc.

Variable aleatoria Binomial

- Es de mucho interés la repetición de experimentos independientes de Bernoulli, por ejemplo n veces.
- En ese caso, el resultado sería una tira de ceros y unos, correspondientes a cada éxito o fracaso:
- Una tal tira sería

$$0010100010$$

  En este caso $n = 10$ y tenemos 3 éxitos y 7 fracasos.
- La variable aleatoria Binomial Y cuenta la cantidad de éxitos en los n experimentos.
- ¿Qué valores puede tomar Y?

- Si no hay ningún éxito tenemos $Y = 0$, si todos son éxitos tenemos $Y = n$.
- Entonces Y toma alguno de los valores $0, 1 \ldots, n$.
- ¿Y con qué probabilidades toma Y cada uno de estos valores?
- Supongamos $n = 2$. ¿Cuántas tiras con 2 ceros o unos tenemos?

- Tenemos 4:

$$00, \quad 01, \quad 10, \quad 11.$$

- Como los resultados son independientes, las probabilidades para estas tiras son:

$$\begin{aligned}
P(00) &= P(0) \times P(0) = (1-p)^2, \\
P(01) &= P(0) \times P(1) = (1-p)p, \\
P(10) &= P(1) \times P(0) = p(1-p), \\
P(11) &= P(1) \times P(1) = p^2,
\end{aligned}$$

- Obtenemos

$$P(Y=0)=(1-p)^2,\; P(Y=1)=2p(1-p),\; P(Y=2)=p^2.$$

- Además

$$\begin{aligned} &P(Y=0)+P(Y=1)+P(Y=2) \\ &\qquad =(1-p)^2+2p(1-p)+p^2=(1-p+p)^2=1. \end{aligned}$$

Veamos el caso $n = 3$

- Cuántas tiras tenemos:

$$\_\ \_\ \_$$

- En este caso $n = 3$ y $2^3 = 8$.
- Contamos la mala suerte pura (todos fracasos)

$$P(Y = 0) = (1 - p)^3$$

- Contamos la buena suerte pura (todos éxitos)

$$P(Y = 3) = p^3$$

- ¿Cómo son los casos intermedios?
- Por ejemplo, para la tira 100 tenemos

$$P(100) = P(1)P(0)^2 = p(1-p)^2.$$

- ¿Cuántas tiras tienen un solo éxito?

$$\underline{1}\ \underline{\ }\ \underline{\ }$$

- Son $3 = \binom{3}{1}$. Entonces

$$P(Y=1) = 3p(1-p)^2 = \binom{3}{1}p(1-p)^2.$$

- Análogamente, dos éxitos es un fracaso, entonces

$$P(Y=2)=3p^2(1-p)=\binom{3}{1}p^2(1-p).$$

- Y se verifica

$$P(Y=0)+P(Y=1)+P(Y=2)+P(Y=3)=$$
$$=(1-p)^3+3(1-p)^2p+3(1-p)p^2+p^3=(1-p+p)^3=1.$$

- Escrito con sumatoria:

$$\sum_{k=0}^{3} P(Y=k)=\sum_{k=0}^{3}\binom{3}{k}(1-p)^{3-k}p^k=1.$$

¿Nos animamos al caso general?

- Tenemos 2^n tiras de ceros y unos
- Si tenemos k éxitos ($0 \leq k \leq n$) los podemos ubicar en n lugares que elegimos de $\binom{n}{k}$ formas diferentes.
- Cada una de esas formas me da un caso de $Y = k$.
- La probabilidad de cada forma es $p^k(1-p)^{n-k}$
- Entonces

$$\boxed{P(Y=k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n}$$

- La variable aleatoria Y se llama binomial, con parámetros n, p.

- Veamos que el modelo está correctamente definido, es decir, la suma de todos los resultados posibles tiene probabilidad uno.
- Para eso tenemos que sumar:

$$\sum_{k=0}^{n} P(Y=k) = \sum_{k=0}^{n} \binom{n}{k} (1-p)^{n-k} p^k = (p+1-p)^n = 1.$$

Ejemplo

En la generala se tiran 5 dados.

(a) Calcular la probabilidad de obtener dos dados con 6.

(b) Calcular la probabilidad de sacar al menos un 6.

Solución:

- El éxito es sacar un 6 al tirar el dado.
- La probabilidad de éxito es $p = 1/6$.
- Y es una binomial con parámetros $n = 5$ y $p = 1/6$
  Tenemos

$$P(Y = 2) = \binom{5}{2} p^2 (1-p)^3 = 10 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 = 0{,}16.$$

- Para la parte (b) tenemos que calcular

$$P(Y \geq 1)$$

- Tenemos

$$P(Y \geq 1) = P(Y = 1) + P(Y = 2) + P(Y = 3) \\ + P(Y = 4) + P(Y = 5)$$

- Es más sencillo:

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1 - p)^5 = 1 - \left(\frac{5}{6}\right)^5 = 0{,}60.$$

Paseo al azar, triangulo de Pascal y variable aleatoria Binomial

- Consideremos una partícula que se desplaza, eligiendo con probabilidad p una unidad para arriba (éxito) o para abajo en cada instante de tiempo (fracaso):
- Llamamos paseo al azar a cada posible trayectoria.

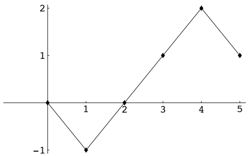


Figura: ¿Qué probabilidad tiene esta trayectoria?

$$P(01110) = p^3(1-p)^2.$$

- ¿Cuál es la probabilidad de que llegue (por cualquier camino) a 1?

$$P(Y = 3) = \binom{5}{3} p^3 (1 - p)^2.$$

Nos interesa calcular la probabilidad de llegar a una altura dada en una cantidad de pasos n

- Si un camino tiene la misma cantidad de ceros y unos, llegará a la misma altura.

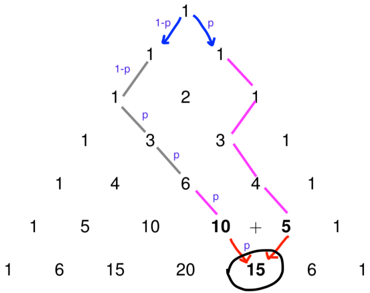
- Nuestro problema es contar de cuantas formas distintas se pueden disponer k unos en n lugares:

$$\underbrace{\underline{0}\ \underline{0}\ \underline{1}\ \underline{0}\ \underline{1}\ \underline{0}\ \underline{1}}_{7 \text{ lugares}}$$

$$C^7_3 = \frac{7!}{3!\,4!} = 35$$

- Y calcular la probabilidad indvidual de esos caminos, que es $p^3(1-p)^4$

Si representamos estas cantidades en una tabla obtenemos el triángulo de Pascal, al que le agregamos probabilidades:



La probabilidad de llegar a ese lugar consiste en tener 4 éxitos y 2 fracasos:

$$P(Y = 4) = \binom{6}{4} p^4 (1 - p)^2 = 15 p^4 (1 - p)^2.$$

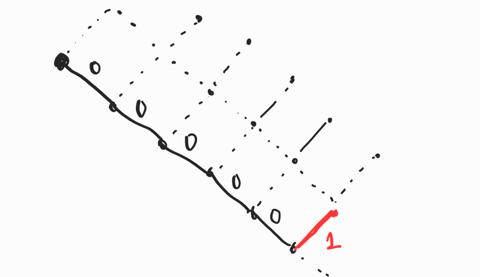
Variable geométrica (la mala suerte no es verdad)

- Se trata de realizar experimentos de Bernoulli hasta el primer éxito.
- El número Y de fracasos necesarios hasta el primer éxito es (por definición) una variable aleatoria geométrica
- Si soy muy afortunado, mi primer experimento es un éxito, me da $Y = 0$
- Si tengo un fracaso y un éxito, tengo $Y = 1$.
- Puede ocurrir que tenga una cantidad de fracasos arbitrariamente grande.
- Entonces Y toma todos los valores naturales:

$$0, 1, \ldots, n, \ldots$$

¿Y cuales son las probabilidades?

- Tenemos: $P(Y = 0) = P(1) = p$.
- Si fracaso una sóla vez: $P(Y = 1) = P(01) = (1 - p)p$.
- Si fracaso dos veces: $P(Y = 2) = P(001) = (1 - p)^2 p$.
- Si fracaso k veces: $P(Y = k) = P(\underbrace{0 \dots 0}_{k} 1) = (1 - p)^k p$.
- Son las trayectorias de la siguiente forma:



- Veamos que en algún momento tenemos un éxito (aunque p sea muy pequeña):

$$\begin{aligned} P(Y=0)+P(Y=1)+\cdots+P(Y=n)+\cdots \\ =p+(1-p)p+\cdots+(1-p)^n p+\cdots \\ =p\times\left[1+(1-p)+\cdots+(1-p)^n+\cdots\right] \end{aligned}$$

- Y la expresión en azul es una serie geométrica de razón $r=1-p$. Tenemos

$$1+r+r^2+\cdots+r^n+\cdots=\frac{1}{1-r}$$

- Entonces

$$
\begin{aligned}
P(Y=0)+P(Y=1)+\cdots+P(Y=n)+\cdots \\
= p \times \left[1+(1-p)+\cdots+(1-p)^n+\cdots\right] \\
= p\frac{1}{1-(1-p)} = 1.
\end{aligned}
$$

- En algún momento tenemos suerte.