Enunciado

Sean X e Y variables aleatorias tales que:

a) Cada variable toma exactamente dos valores.

b) $E(XY) = E(X)E(Y)$

Se pide demostrar que X e Y son independientes

Solución

Sean X e Y las variables aleatorias. Del apartado a del enunciado deducimos que son variables aleatorias discretas y, por lo tanto, podemos definir su soporte:

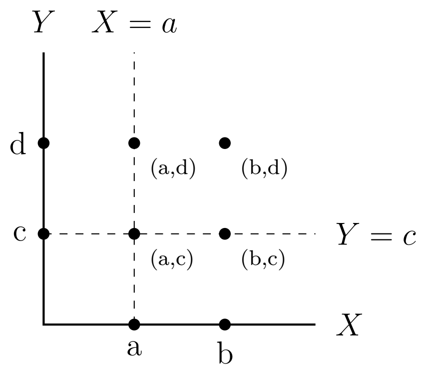
$$S_X = \{a, b\} \qquad S_Y = \{c, d\}$$

Con el soporte definido, definimos la función de masa de probabilidad para cada una variable:

$$p_x(t) = \begin{cases} p & \text{si } t = a \\ 1-p & \text{si } t = b \end{cases} \qquad p_y(t) = \begin{cases} q & \text{si } t = c \\ 1-q & \text{si } t = d \end{cases}$$

Del mismo modo, ya podemos definir el soporte del vector aleatorio (X, Y)

$$S_{(X,Y)} = \Big\{(a,c),\ (a,d),\ (b,c),\ (b,d)\Big\}$$



Para demostrar la independencia de las variables aleatorias X e Y, vamos a aplicar:

$$X, Y \text{ discretas, son independientes} \iff p_{(X,Y)}(t_1, t_2) = p_X(t_1) \cdot p_Y(t_2) \ \ \forall (t_1, t_2) \in S_x \times S_y$$

Vamos ahora a estudiar la probabilidad de que el vector (X, Y) tome cada valor en su soporte. En concreto, vamos a intentar dejar todas las probabilidades en función de $p_{(X,Y)}(a, c)$, a la que denominaremos $r = p_{(X,Y)}(a, c)$

Lo haremos primero para los valores (a, d) y (b, c). Para estos, podemos aprovechar que solo hay dos elementos en $S_{(X,Y)}$ para los que X o Y tomen un determinado valor. De ahí podemos obtener las siguientes relaciones:

$$\begin{aligned} p_X(a) = p_{(X,Y)}(a,c) + p_{(X,Y)}(a,d) \implies p_{(X,Y)}(a,d) &= p_X(a) - p_{(X,Y)}(a,c) = p - p_{(X,Y)}(a,c) \\ &= p - r \\ p_Y(c) = p_{(X,Y)}(a,c) + p_{(X,Y)}(b,c) \implies p_{(X,Y)}(b,c) &= p_Y(c) - p_{(X,Y)}(a,c) = q - p_{(X,Y)}(a,c) \\ &= q - r \end{aligned}$$

La probabilidad del último punto la podemos obtener aplicando la fórmula de la probabilidad total:

$$\begin{aligned}
p_{(X,Y)}(b,d) &= 1 - \Big(p_{(X,Y)}(a,c) + p_{(X,Y)}(b,c) + p_{(X,Y)}(a,d)\Big) \\
&= 1 - \Big(p_{(X,Y)}(a,c) + p - p_{(X,Y)}(a,c) + q - p_{(X,Y)}(a,c)\Big) \\
&= 1 - \Big(p + q - p_{(X,Y)}(a,c)\Big) = 1 - p - q + p_{(X,Y)}(a,c) \\
&= 1 - p - q + r
\end{aligned}$$

Con toda esta información sobre las probabilidades en cuenta, podemos ahora calcular las esperanzas.

$$\begin{aligned}
E(X) &= \sum_{t\in S_X} t \cdot p_X(t) = p \cdot a + (1-p) \cdot b \\
E(Y) &= \sum_{t\in S_Y} t \cdot p_Y(t) = q \cdot c + (1-q) \cdot d
\end{aligned}$$

Para calcular la esperanza de XY, vamos a utilizar la siguiente fórmula, donde $\Phi(u,v) = uv$.

$$E(XY) = E(\Phi(X,Y)) = \sum_{(t_1,t_2)\in S_x\times S_y} \Phi(t_1,t_2)\cdot p_{(X,Y)}(t_1,t_2) = \sum_{(t_1,t_2)\in S_x\times S_y} t_1 t_2 \cdot p_{(X,Y)}(t_1,t_2)$$

Si desarrollamos el último sumatorio y sustituímos las probabilidades:

$$\begin{aligned}
E(XY) &= ac \cdot p_{(X,Y)}(a,c) + ad \cdot p_{(X,Y)}(a,d) + bc \cdot p_{(X,Y)}(b,c) + bd \cdot p_{(X,Y)}(b,d) \\
&= ac \cdot r + ad \cdot \Big(p - r\Big) + bc \cdot \Big(q - r\Big) + bd \cdot \Big(1 - p - q + r\Big)
\end{aligned}$$

Por otra parte, podemos calcular el producto $E(X) \cdot E(Y)$:

$$\begin{aligned}
E(X) \cdot E(Y) &= ac \cdot pq + ad \cdot p(1-1) + bc \cdot q(1-p) + bd \cdot (1-p)(1-q) = \\
&= ac \cdot pq + ad \cdot (p - pq) + bc \cdot (q - pq) + bd \cdot (1 - p - q + pq)
\end{aligned}$$

Ahora podemos comparar los resultados que hemos obtenido para $E(X,Y)$ y para el producto $E(X)\cdot E(Y)$. Si los igualamos podemos ver que, necesariamente, $r = pq$ ya que esta es la única incógnita que tenemos en la primera ecuación. A partir de aquí podemos comprobar la condición que habíamos pedido inicialmente: $\forall i,j \quad p_{(X,Y)}(i,j) = p_X(i) \cdot p_Y(j)$.

$$\begin{aligned}
p_{(X,Y)}(a,c) &= pq & &= p_X(a) \cdot p_Y(c) \\
p_{(X,Y)}(a,d) &= p - pq &= p(1-q) &= p_X(a) \cdot p_Y(d) \\
p_{(X,Y)}(b,c) &= q - pq &= (1-p)q &= p_X(b) \cdot p_Y(c) \\
p_{(X,Y)}(b,d) &= 1 - p - q + pq &= (1-p)(1-q) &= p_X(b) \cdot p_Y(d)
\end{aligned}$$

■