1. Variables aleatorias discretas

Definición: Una variable aleatoria Y se denomina discreta si puede adoptar solo una cantidad finita o infinita contable de valores distintos.

Por ejemplo una variable de interés puede ser cuantos profesores del Departamento de Computo Científico y Estadística están dando clases en horario 3-4 los lunes. Claramente esa variable es discreta, ya que la cantidad de profesores es una cantidad finita.

Definición: La probabilidad de que Y adopte el valor y, $P(Y = y)$, se define como la suma de las probabilidades de los puntos muestrales de S que tienen asignado el valor y. A veces se representará $P(Y = y)$ como $p(y)$.

Definición: La distribución de probabilidad para una variable discreta Y puede representar mediante una fórmula, tabla o gráfica, que proporciona $p(y) = P(Y = y)$ para toda y.

Ejemplo. El ministerio de salud analizó los pozos de un país para detectar dos clases de impurezas (A y B) que, por lo común, contiene el agua potable. En 20 % de ellos no se encontró ninguna de las impurezas. En 40 % se detecto la impureza A y en 50 % la impureza B. (Es obvio que algunos contenían los dos tipos). Encuentre la distribución de probabilidad de Y, es decir, la cantidad de impurezas que contiene un pozo elegido al azar.

Solución.
Como la variable aleatoria Y cuenta la cantidad de impurezas que puede tener un pozo, entonces Y tiene tres valores posibles:

0 en el caso que no tenga ninguna impureza.
1 en el caso que tenga la impureza A o la impureza B.
2 en el caso que tenga tanto la impureza A como la B.

Es claro que

$$p(0) = P(Y = 0) = 0{,}2$$

Ahora, nos dan los porcentajes de pozos donde se consiguen las impurezas A y B individualmente. Es claro que si el 20 % no tiene impurezas, entonces el 80 % si tienen al menos una impureza (ya que es el complemento del conjunto). Como las impurezas A y B, individualmente, suman 90 %, entonces ese 10 % sobrante representa la intersección de ambos conjuntos, luego:

$$p(2) = P(Y = 2) = 0{,}1$$

Luego, como la suma de las probabilidades de todo el conjunto debe ser 1, entonces

$$p(1) = P(Y = 1) = 0{,}7$$

Entonces

$$P(Y=y)=\begin{cases} 0{,}2 & \text{si} \quad Y=0 \\ 0{,}7 & \text{si} \quad Y=1 \\ 0{,}1 & \text{si} \quad Y=2 \end{cases}$$

<u>Ejemplo</u>. De las personas que llegan a un banco de sangre, 1 de cada 3 tiene tipo de sangre O^+, y 1 de cada 15 tiene sangre tipo O^-. Supongamos que se eligen aleatoriamente a 3 donadores. Si X es la cantidad de donadores con sangre tipo O^+ y Y la de los que tienen tipo O^-, encuentre las distribuciones de probabilidad de X y Y. Encuentre también la distribución de probabilidad para $X+Y$, el número de donadores con sangre tipo O.

<u>Solución</u>.
Comencemos viendo que la probabilidad de que una persona que llega al banco de sangre tenga del tipo O^+ es $1/3$. La variable aleatoria X puede tener los valores 0, 1, 2 y 3; que representa la cantidad de donantes que llegan al banco con tipo de sangre O^+. La probabilidad de que un paciente no sea O^+ es $2/3$ (que es el complemento de la probabilidad de ser O^+). La probabilidad de que ningún paciente tenga sangre O^+, es:

$$P(X=0)=\frac{2}{3}\frac{2}{3}\frac{2}{3}=\frac{8}{27}.$$

Luego, para calcular la probabilidad de que llegue un donante con tipo de sangre O^+, haremos:

$$\begin{aligned} P(X=1) &= \binom{3}{1}\frac{1}{3}\frac{2}{3}\frac{2}{3}=\frac{3!}{2!1!}\frac{4}{27} \\ &= 3\frac{4}{27}=\frac{12}{27}. \end{aligned}$$

Es de hacer notar que lo que se hace es, escoger uno de los tres donantes, que sera el que tiene tipo O^+, por ello se coloca el combinatorio de 3 en 1. La probabilidad de esa persona es $1/3$, y las dos restantes tendrán probabilidad de $2/3$. Pensando de la misma manera:

$$P(X=2)=\binom{3}{2}\frac{1}{3}\frac{1}{3}\frac{2}{3}=\frac{6}{27}.$$

Finalmente:

$$P(X=3)=\frac{1}{3}\frac{1}{3}\frac{1}{3}=\frac{1}{27}.$$

Entonces la distribución de probabilidad de X es:

$$P(X=x)=\begin{cases} 8/27 & \text{si} \quad X=0 \\ 12/27 & \text{si} \quad X=1 \\ 6/27 & \text{si} \quad X=2 \\ 1/27 & \text{si} \quad X=3 \end{cases}$$

Realizamos el mismo proceso para hallar la distribución de Y, tomando en cuenta que la probabilidad de que una persona tenga tipo de sangre O^- es $1/15$ y la probabilidad de que no tenga dicho tipo de sangre es $14/15$.

$$P(Y=0)=\frac{14}{15}\frac{14}{15}\frac{14}{15}=\frac{2744}{3375}.$$

$$P(Y=1)=\binom{3}{1}\frac{1}{15}\frac{14}{15}\frac{14}{15}=\frac{588}{3375}.$$

$$P(Y=2)=\binom{3}{2}\frac{1}{15}\frac{1}{15}\frac{14}{15}=\frac{42}{3375}.$$

$$P(Y=3)=\frac{1}{15}\frac{1}{15}\frac{1}{15}=\frac{1}{3375}.$$

Por lo tanto la distribución de Y es:

$$P(Y=y)=\begin{cases} 2744/3375 & \text{si} \quad Y=0 \\ 588/3375 & \text{si} \quad Y=1 \\ 42/3375 & \text{si} \quad Y=2 \\ 1/3375 & \text{si} \quad Y=3 \end{cases}$$

Supongamos ahora que $Z=X+Y$. Para estudiar esta variable aleatoria, debemos ver cual es la probabilidad de tener tipo de sangre O. Para ello, fijese que la probabilidad de que un donante tenga tipo de sangre O^+ es $1/3$, lo cual también se puede escribir como $5/15$ (ambas fracciones representan el mismo número); mientras que la probabilidad de que un donante tenga sangre de tipo O^- es $1/15$. Por lo tanto podemos decir que la probabilidad de que un donante tenga sangre tipo O es $6/15$. Como Z también puede tomar los valores que X y Y, entonces lo analizaremos de igual forma, tomando en cuenta, que la probabilidad de no tener tipo de sangre O es $9/15$.

$$P(Z=0)=\left(\frac{9}{15}\right)^3=\frac{729}{3375}.$$

$$P(Z=1)=\binom{3}{1}\frac{6}{15}\left(\frac{9}{15}\right)^2=\frac{1458}{3375}.$$

$$P(Z=2)=\binom{3}{2}\left(\frac{6}{15}\right)^2\frac{9}{15}=\frac{972}{3375}.$$

$$P(Z=3)=\left(\frac{6}{15}\right)^3=\frac{216}{3375}.$$

Finalmente la distribución de $X+Y$ es:

$$P(Z = z) = P(X + Y = z) = \begin{cases} 729/3375 & \text{si} \quad X + Y = 0 \\ 1458/3375 & \text{si} \quad X + Y = 1 \\ 972/3375 & \text{si} \quad X + Y = 2 \\ 216/3375 & \text{si} \quad X + Y = 3 \end{cases}$$

<u>Teorema</u>: Cualquier distribución de probabilidades discreta debe satisfacer lo siguiente:

1. $0 \leq p(y) \leq 1$ para todo valor de y.
2. $\sum_y p(y) = 1$.