Ejercicio 1

Una estación de radio tiene dos grupos de oyentes, los jóvenes y los viejos. Se sabe que si el oyente es joven hay una probabilidad del 95% de que le guste el programa 1, una probabilidad del 5% de que le guste el programa 2, una probabilidad del 2% de que le guste programa 3 y una probabilidad del 20% de que le guste el programa 4. Por otro lado, si el oyente es viejo, hay una probabilidad del 3% de que le guste el programa 1, una probabilidad del 82% de que le guste el programa 2, una probabilidad del 34% de que le guste el programa 3 y una probabilidad del 92% de que le guste el programa 4. Se sabe también que el 90% de los oyentes son viejos.

Un nuevo oyente escucha los programas 1 y 3 pero no le gustan los programas 2 y 4. Calcular la probabilidad de que este oyente sea joven y la probabilidad de que sea viejo.

Solución

Sea $Edad$ la variable aleatoria que indica la edad de un oyente, y toma valores V por viejo y J por joven. Tenemos que $P(Edad = V) = 0.9$. Cómo sólo se clasifica a los oyentes en las categorías joven y viejo, entonces $P(Edad = J) = 1 - P(Edad = V) = 0.1$. Llamemos G_i a las variables binarias (con valores "Sí" y "No") indican si a un oyente le gusta el programa i.

La tabla 1 resume las probabilidades condicionales de G_i dada la Edad. Esto se desprende directamente de los datos del enunciado.

Joven		
i	$P(G_i) = Sí$	$P(G_i) = No$
1	*0.95*	0.05
2	*0.05*	0.95
3	*0.02*	0.98
4	*0.2*	0.8

Viejo		
i	$P(G_i) = Sí$	$P(G_i) = No$
1	*0.03*	0.97
2	*0.82*	0.18
3	*0.34*	0.66
4	*0.92*	0.08

Tab. 1: Probabilidades condicionales de que le guste cada programa a un oyente dada su condicion de viejo o joven

Por otro lado, utilizando estos datos y el teorema de la probabilidad total podemos calcular

$$P(G_i) = P(G_i|Edad = J)P(Edad = J) + P(G_i|Edad = V)P(Edad = V)$$

Las distribuciones de G_i se muestran en la tabla 2.

i	$P(G_i) = Sí$	$P(G_i) = No$
1	0.122	0.878
2	0.743	0.257
3	0.308	0.692
4	0.848	0.152

Tab. 2: Distribuciones de las variables G_i

Llamemos $NuevoOyente$ a la condición ($G_1 = $ Sí $\wedge G_2 = No \wedge G_3 = $ Sí $\wedge G_4 = No$).
Queremos calcular $P(Edad = V|NuevoOyente)$ y $P(Edad = J|NuevoOyente)$.

Tenemos

$$P(Edad = V|NuevoOyente) = \frac{P(NuevoOyente|Edad = V) \cdot P(Edad = V)}{P(NuevoOyente)}$$

$$P(Edad = J|NuevoOyente) = \frac{P(NuevoOyente|Edad = J) \cdot P(Edad = J)}{P(NuevoOyente)}$$

Asumiendo que las variables G_i tienen independencia condicional dada $Edad$, se puede calcular $P(NuevoOyente|Edad = X)$ como:
$P(G_1 = Sí|Edad = V) \cdot P(G_2 = No|Edad = V) \cdot P(G_3 = Sí|Edad = V) \cdot P(G_4 = No|Edad = V)$

Y así $P(NuevoOyente|Edad = V) = 0.03 \cdot 0.18 \cdot 0.34 \cdot 0.08 \approx 0.00015$. Con un procedimiento análogo vemos que $P(NuevoOyente|Edad = J) = 0.95 \cdot 0.95 \cdot 0.02 \cdot 0.8 = 0.01444$

Además $P(NuevoOyente)$ se puede calcular, por probabilidad total, como:

$$P(NuevoOyente|Edad = V) \cdot P(Edad = V) + P(NuevoOyente|Edad = J) \cdot P(Edad = J)$$

Así $P(NuevoOyente) = 0.00015 \cdot 0.9 + 0.01444 \cdot 0.1 \approx 0.00158$

Finalmente tenemos

$$P(Edad = V|NuevoOyente) = \frac{0.00015 \cdot 0.9}{0.00158} \approx 0.09$$

y de la misma forma:

$$P(Edad = J|NuevoOyente) = \frac{0.01444 \cdot 0.1}{0.00158} \approx 0.91$$

.

Observación

Asumir la independencia condicional de las G_i dada la edad hace que podamos representar el dominio mediante la red Bayesiana que muestra la figura 1. De hecho se puede notar que calculamos todas las probabilidades de este problema a partir de las probabilidades de $Edad$ y de las variables condicionadas $G_i|Edad$, que son exactamente $G_i|Pa(G_i)$.

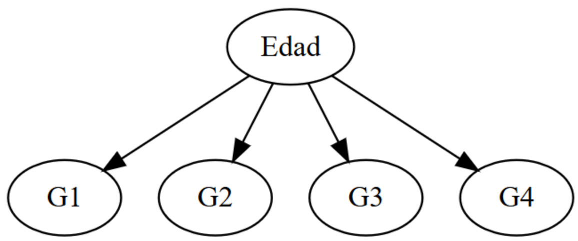

Fig. 1: Red bayesiana, asumiendo independencia condicional de las G_i dada Edad