Serie 1

· $U \perp X$

Afirmación 1: $U \perp X$

- Camino entre U y X : $U \leftarrow T \rightarrow X$

Este es un Common Cause ($A \leftarrow B \rightarrow C$) con $B=T$.

  - Si T no está observado, el Common Cause es activo.
  - Si T está observado, el Common Cause es inactivo.
- Condición: No hay ninguna variable condicional, así que T no está observado.
- Resultado: El camino es activo, por lo tanto, U y X no son independientes .

Falsa

· $U \perp X|T$

Afirmación 2: $U \perp X|T$

- Camino entre U y X : $U \leftarrow T \rightarrow X$

Este es un Common Cause ($A \leftarrow B \rightarrow C$) con $B=T$.

  - Si T está observado, el Common Cause es inactivo .
- Condición: Aquí, T está condicionalmente observado.
  - Resultado: El camino es inactivo, por lo tanto, U y X son independientes condicionalmente dado T .

Verdadera

· $V \perp W|Y$

Afirmación 3: *V⊥W |Y*

- Camino entre *V* y *W*: *V→Y←W*

Este es un Common Effect (*A→B←C*) con *B=Y*.

    - Si *Y* está observado, el Common Effect es activo.
    - Si *Y* no está observado, el Common Effect es inactivo.
- Condición: Aquí, *Y* está condicionalmente observado.
- Resultado: El camino es activo, por lo tanto, *V* y *W* no son independientes condicionalmente dado *Y*.

Falsa

· $V \perp W|T$

Afirmación 4: V⊥W|T

· Camino entre V y W: V→Y←W

Este es un Common Effect (A→B←C) con B=Y.

- Si Y está observado, el Common Effect es activo.
- Si Y no está observado, el Common Effect es inactivo.

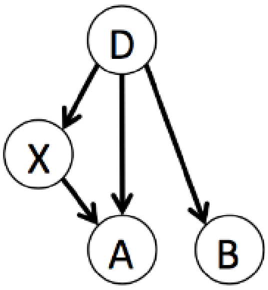
· Condición: Aquí, T está condicionalmente observado, pero Y no está observado.

· Resultado: El camino es inactivo, por lo tanto, V y W son independientes condicionalmente dado T.

Verdadera

Serie 2

$P(A\|D,X)$			
$+d$	$+x$	$+a$	0.9
$+d$	$+x$	$-a$	0.1
$+d$	$-x$	$+a$	0.8
$+d$	$-x$	$-a$	0.2
$-d$	$+x$	$+a$	0.6
$-d$	$+x$	$-a$	0.4
$-d$	$-x$	$+a$	0.1
$-d$	$-x$	$-a$	0.9



$P(D)$	
$+d$	0.1
$-d$	0.9

$P(X\|D)$		
$+d$	$+x$	0.7
$+d$	$-x$	0.3
$-d$	$+x$	0.8
$-d$	$-x$	0.2

$P(B\|D)$		
$+d$	$+b$	0.7
$+d$	$-b$	0.3
$-d$	$+b$	0.5
$-d$	$-b$	0.5

a) P(+d,+a)

Variables de Consulta: D, A
Ocultas: X
Evidencia: No hay

Eliminación de Variables: Eliminar X
Tablas a trabajar:
P(D)   P(X|D)   P(A|D,X)

Sabemos según la pregunta D es +d y A es +a por tanto podemos reducir las tablas a:
P(+d)   P(X|+d)   P(+a| +d,X)

P(+d) = 0.1

P(X|+d)

D	X	P(X|D)
+d	+x	0.7
+d	-x	0.3

P(+a|+d,X)

D	X	A	P(X|D)
+d	+x	+a	0.9
+d	-x	+a	0.8

Join: P(X|+d)*P(+a|+d,X)

(0.7*09)+(0.3*0.8)=0.63+0.24 =0.87

Sum(eliminación): P(+d)*P(+a,X,+d) = 0.1*0.87 = **0.087**

b) **P(+d|+a)** = $\frac{P(+d,+a)}{P(+a)}$

Se obtuvo **P(+d,+a)** del inciso anterior o sea 0.087

P(+a) = Σ P(+a,D) = P(+a,+d)+P(+a,-d)

P(+a,+d) es lo mismo que **P(+d,+a)** o sea 0.087, por tanto solo debemos encontrar P(+a,-d)

P(+a,-d) = Σ P(+a,X,-d) entonces eliminación de variables, para eliminar X

Σ P(+a,X,-d) = P(+a | -d, X) P(X| -d) P(-d)

D	X	A	P(X|D)
-d	+x	+a	0.6
-d	-x	+a	0.1

D	X	P(X|D)
-d	+x	0.8
-d	-x	0.2

0.9*(0.6*08)+(0.2*0.1)=0.45

Entonces P(+d|+a) = $\frac{P(+d,+a)}{P(+a)} = \frac{P(+d,+a)}{P(+a,+d)+P(+a,-d)} = \frac{0.087}{0.087+0.45} = 0.162$

c) **P(+d|+b)** = $\frac{P(+d,+b)}{P(+b)} = \frac{P(+b|+d)\,P(+d)}{P(+b)}$

P(+b|+d) lo conocemos de la tabla P(B|D) y tambien P(+d) de P(D)
por tanto solo queda calcular P(+b) = Σ P(+b,D) = P(+b,+d) + P(+b,-d)
P(+b,+d) + P(+b,-d) = P(-d)P(+b|-d) + P(+d)P(+b|+d)

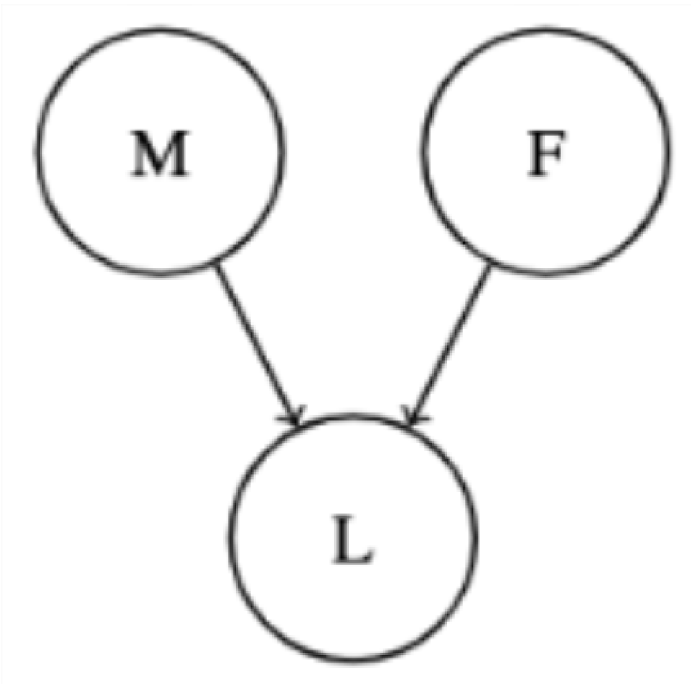
Según las tablas
P(-d) = 0.9
P(+b|-d) = 0.5
P(+d) = 0.1
P(+b|+d) = 0.7

P(+b) = Σ P(+b,D) = (0.9*0.5)+(0.1*0.7) = 0.52

Sabiendo que P(+b|+d) = 0.7 y P(+d) = 0.1

Entonces P(+d|+b) $= \frac{0.7*0.1}{0.52} =$ 0.135

Serie 3



M	F	$P(L=+l\mid M,F)$
$+m$	$+f$	0.05
$+m$	$-f$	0.01
$-m$	$+f$	0.9
$-m$	$-f$	0.02

Valor de *M*	**Símbolo**	**Probabilidad *P*(*M*)**
"Tiene clase matutina"	$+m$	0.7
"No tiene clase matutina"	$-m$	1 - 0.7 = **0.3**

Valor de *F*	**Símbolo**	**Probabilidad *P*(*F*)**
"Amigos regresan"	$+f$	0.8
"Amigos no regresan"	$-f$	1 - 0.8 = **0.2**

a) Fórmula para la distribución de probabilidad conjunta

La red Bayesiana tiene la siguiente estructura:

- M (clase matutina) y F (amigos regresan) son nodos padres de L (disfrute universitario)
- La distribución de probabilidad conjunta en una red bayesiana se expresa como:

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P\left(X_i \mid Padres(X_i)\right)$$

Respuesta:

$$P(M, F, L) = P(M) \times P(F) \times P(L \mid M, F)$$

Donde:

- $P(M)$ es la probabilidad marginal de tener clase matutina
- $P(F)$ es la probabilidad marginal de que los amigos regresen
- $P(L|M,F)$ es la probabilidad condicional de disfrutar dado M y F (tabla proporcionada)

b) Probabilidad de que sus amigos regresen y no tenga clase matutina

Queremos calcular $P(F=+f, M=-m)$:

Datos:

- $P(F=+f)=0.8$
- $P(M=+m)=0.7$, por lo tanto $P(M=-m)=1-0.7=0.3$

Como M y F son independientes (no tienen padres en común en la red):

$$P(F = +f, M = -m) = P(F = +f) \times P(M = -m) = 0.8 \times 0.3 = 0.24$$

Respuesta: La probabilidad es 0.24 o 24%.

c) Probabilidad de que sus amigos regresen dado que le está gustando el semestre

Queremos calcular $P(F{=}{+}f|L{=}{+}l)$. Usaremos el teorema de Bayes:

$$P(F=+f \mid L=+l) = \frac{P(L=+l \mid F=+f) \times P(F=+f)}{P(L=+l)}$$

1. Calcular $P(L{=}{+}l \mid F{=}{+}f)$

Marginalizamos sobre M:

$$P(L=+l \mid F=+f) = \sum_{M} P(L=+l \mid M, F=+f)P(M)$$

$$=P(L=+l \mid M=+m,F=+f)P(M=+m)+P(L=+l \mid M=-m,F=+f)P(M=-m)$$

$$=0.05\times0.7+0.9\times0.3$$

$$=0.035+0.27=0.305$$

2. Calcular P(L=+l)$P(L{=}{+}l)$

Marginalizamos sobre M y F:

$$P(L=+l) = \sum_{M,F} P(L=+l \mid M,F)P(M)P(F)$$

$$=P(L=+l|M=+m,F=+f)P(M=+m)P(F=+f)+$$

$$P(L=+l|M=+m,F=-f)P(M=+m)P(F=-f)+$$

$$P(L=+l|M=-m,F=+f)P(M=-m)P(F=+f)+$$

$$P(L=+l|M=-m,F=-f)P(M=-m)P(F=-f)$$

$$=0.05\times0.7\times0.8+0.01\times0.7\times0.2+0.9\times0.3\times0.8+0.02\times0.3\times0.2$$

$$=0.028+0.0014+0.216+0.0012$$

$$=0.2466$$

3. Aplicar Bayes

$$P(F=+f \mid L=+l) = \frac{0.305 \times 0.8}{0.2466}$$

$$= \frac{0.244}{0.2466} \approx 0.9895$$

Respuesta: La probabilidad es aproximadamente 0.9895 o 98.95%.

Serie 4

1. Calcular los log-odds (z) para cada dato

La fórmula de los log-odds es:

$$z = \beta_0 + \beta_1 \cdot X$$

Donde:

- $\beta 0$=−58.0264 (intercepto)
- $\beta 1$=0.0116 (coeficiente de ingresos)

Aplicamos para cada cliente:

Ingresos (X)	**Cálculo de *z***	**Log-odds (z)**
2000	−58.0264+0.0116·2000	−58.0264+23.2=−34.8264
4000	−58.0264+0.0116·4000	−58.0264+46.4=−11.6264
6000	−58.0264+0.0116·6000	−58.0264+69.6=11.5736
8000	−58.0264+0.0116·8000	−58.0264+92.8=34.7736

---

2. Aplicar la función sigmoide para obtener probabilidades

La sigmoide convierte log-odds (*z*) en probabilidades (*P*):

$$P = \frac{1}{1 + e^{-z}}$$

Calculamos para cada *z*:

Ingresos (X)	**Log-odds (z)**	**Cálculo de *P***	**Probabilidad *P***
2000	−34.8264	$\frac{1}{1 + e^{34.8264}}$	≈0 (prácticamente 0)

4000	−11.6264	$\frac{1}{1+e^{11.6264}}$	≈0 (muy cercano a 0)
6000	11.5736	$\frac{1}{1+e^{-11.5736}}$	≈1 (muy cercano a 1)
8000	34.7736	$\frac{1}{1+e^{-34.7736}}$	≈1 (prácticamente 1)

Nota:

- Si z es negativo grande, $P \approx 0$.
- Si z es positivo grande, $P \approx 1$.

---

3. Predecir para nuevos valores de *X*

Usamos el mismo proceso para $X = \{3000, 5000, 7000\}$, con umbral 0.5:

Ingresos (X)	Probabilidad *P*	Predicción
3000	≈0	0 (no compra)
5000	0.493	0 (no compra)
7000	≈1	1 (compra)

Interpretación

- **El modelo "aprendió" que**:
  - Clientes con ingresos **altos** ($X \geq 6000$) tienden a comprar ($y=1$).
  - Clientes con ingresos **bajos** ($X \leq 5000$) no compran ($y=0$).
- **El punto de corte está cerca de $X=5000$**:
  - En $X=5000$, $P \approx 0.5$, lo que indica incertidumbre.

Serie 5

Calcular las métricas clave:

Precisión para spam y Recall para spam. El objetivo es encontrar el modelo que:

- Maximice la precisión para spam (evitar falsos positivos, FP).
- Mantenga un buen recall (capturar la mayoría del spam, TP).

Métricas:

- **Precisión (Precision)**: Cuando el modelo predice la clase positiva, ¿cuántas veces acierta? = TP / (TP + FP)
- **Sensibilidad (Recall)**: Proporción de ejemplos positivos correctamente clasificados.
    - sensibilidad = TP / (TP + FN)

Modelo	Precisión (spam)	Recall (spam)	Resultado
A	58.3%	70%	**Mejor equilibrio**: Precisión aceptable y recall decente.
B	31.0%	90%	**Más spam detectado (TP)**, pero muchos FP (200).
C	80%	20%	Alta precisión, pero recall muy bajo (pierde 80% del spam).
D	14.3%	50%	El peor: baja precisión y recall mediocre.

- a): El Modelo A es el mejor equilibrado.

Para resolver la pregunta b), en la tabla anterior vemos que el modelo B es la respuesta, porque prioriza la cantidad de spam detectado (recall alto), incluso si eso implica clasificar incorrectamente muchos no-spam como spam (FP altos), o de otra manera:

Modelo	TP (spam detectado)	FP (errores graves)	Recall	Precisión
A	70	50	70%	58.3%
B	**90** (más alto)	**200** (alto)	90%	31.0%
C	20 (bajo)	**5** (más bajo)	20%	80%
D	50	300 (más alto)	50%	14.3%

- b): El Modelo B detecta más spam (TP alto) pero con muchos errores (FP alto).

Nota: Naive Bayes asume independencia entre características, lo que puede explicar los altos FP en algunos modelos (ej. B).