Práctica 5
Variables aleatorias conjuntas

1. Ejercicio 1

El 60 % de los aspirantes a un cargo ejecutivo de mucha responsabilidad tienen el equilibrio emocional como para sobrellevar situaciones de mucho estrés. Clasificá como "adecuados" a aquellos con el adecuado equilibrio emocional. Como parte del proceso de selección el aspirante debe rendir un exámen que consta de 5 preguntas sobre cómo gestionar distintas situaciones de conflicto. Suponé que la probabilidad de que un aspirante adecuado conteste satisfactoriamente a una pregunta es 0,8 y la probabilidad de que un aspirante inadecuado conteste satisfactoriamente a una pregunta es 0.5. Suponé además que para una persona dada, las respuestas son independientes entre sí. Definí las siguientes variables aleatorias. X : la variable Bernoulli que vale 1 si un aspirante es adecuado y 0 si no lo es, Y : número de respuestas satisfactorias.

(a) Calulá la distribución conjunta de X e Y

(b) Calculá la distribución marginal de Y

(c) Calculá la distribución condicional de X dado $Y = y$ para todo $y \in \{0, 1, 2, 3, 4, 5\}$

(d) Si un aspirante hubiese salido airoso de todas las restantes intancias del proceso de selección, y en el exámen hubiese respondido cuatro preguntas adecuadamente, ¿le ofrecerías el cargo? Y si hubiese respondido solo dos preguntas adecuadamente, ¿se lo ofrecerías? Justificá tu respuesta.

2. Ejercicio 2

Suponé que de los libros antiguos que hay en el mundo, el 10 % son raros. Suponé también que de los libros antiguos con tapa dorada, también el 10 % son raros. Suponé además que el 30 % de los libros antiguos tienen tapa dorada. Te enteraste que tengo un libro antiguo. Llama X a la variable aleatoria Bernoulli que vale 1 si mi libro es raro y llamá Y a la variable aleatoria Bernoulli que vale 1 si mi libro antiguo tiene tapa dorada. Calculá la distribución conjunta de X e Y y demostrá que X e Y son independientes. Suponé ahora que te cuento que vendí mi libro a un librería de libros antiguos que exhibe un libro en la vidriera si y solo si el libro es raro o tiene tapa dorada. Llamá Z a la variable aleatoria Bernoulli que vale 1 si mi libro será exhibido en la vidriera y vale 0 de otro modo. Calculá la probabilidad de éxito de Z. ¿Son X e Y independientes dado $Z = 1$?

Nota: este problema ilustra el así llamado "*sesgo de Berkson*"que en algunas disciplinas también lo denominan "*sesgo de selección*". Este sesgo ocurre cuando dos variables que no tienen relación entre sí (en nuestro ejemplo X e Y) se vuelven dependientes si nos enteramos - es decir, al condicionar en - otra variable aleatoria (en nuestro ejemplo, Z).

3. Ejercicio 3

El número de emergencias médicas que sufre un alumno de una cierta universidad durante el período lectivo es una v.a. Pois(λ). Supongamos que hay dos tipos de alumnos, aquellos mas propensos a sufrir emergencias médicas, para los cuales $\lambda = 2$ y aquellos poco propensos a las emergencias médicas, para las cuales $\lambda = 0{,}2$. El 10 % de los alumnos son del tipo de los propensos a sufrir emergencias médicas. Si cada asistencia a una emergencia médica tiene un costo para la universidad de 2000 pesos, ¿cuál es el gasto esperado en asistencia medica para cada alumno por ciclo lectivo? ¿Cuál es el desvío estandard del gasto por alumno y período? ¿Cuál es la probabilidad de que un nuevo ingresante no ocasione ningún gasto en emergencias médicas durante el próximo ciclo lectivo? ¿Cuál es la probabilidad de que ocasione exactamente 4000 pesos en gastos por emergencias médicas? Para resolver este problema, defini claramente las variables aleatorias que usas.

4. Ejercicio 4

El número de clientes que entran a un comercio de prendas de vestir en un día dado es una variable aleatoria $Pois\,(30)$. Suponé que la cantidad de prendas que compra un cliente en una visita al comercio es una variable aleatoria Poiss(2) y que las decisiones de comprar o no una prenda de cada cliente no está influida por las compras de otros clientes ni por cuán lleno está el comercio.

(a) ¿Cuáles son la esperanza y el desvío estandard del número de prendas vendidas en un día por el comercio?.

(b) Si la ganancia neta de una compra es una variable aleatoria con media 100 y desvío estandard 20 y es independiente del número de compras y de visitas al comercio ¿cuáles son la media y el desvío estandard de la ganancia neta diaria del comercio?

5. Ejercicio 5

El siguiente juego es una variación del juego que se conoce en economía como el juego de las billeteras. Supongamos que hay n personas no relacionadas entre sí, la persona i tiene en su billetera una cantidad X_i de pesos, la cual es desconocida para el resto del grupo. Se subasta la suma $V = X_1 + ... + X_n$ del dinero en las n billeteras, cada una de las n personas hará una oferta y ganará la subasta aquel cuya oferta sea la más alta. Lo único que es sabido por todas las personas es que cada billetera puede a lo sumo tener 100 pesos.

(a) Si la persona i, elegida al azar, NO conoce cuanto dinero hay en su billetera, cuál es la esperanza V?

(b) Si la persona i, elegida al azar, sólo conoce la cantidad de dinero que hay en su billetera, cuál es la esperanza de V ahora?

(c) Suponé que cada persona i oferta el valor esperado en el punto (b). Suponé que la persona $i = 1$ gana la subasta. Condicional en conocer el valor de X_1 y en haber ganado la subasta, cuál es la esperanza de V? Explicá intuitivamente por qué esta esperanza es menor que la que encontraste en el punto (b).

6. Ejercicio 6

Vas a invertir 10000 \$ que tenés la opción de distribuidos en en tres portafolios. Llamá X_1, X_2 y X_3 a las tasas de retorno al cabo de un año de los tres portafolios expresadas en proporciones -NO en porcentajes-, de modo que por ejemplo, si invertís 3000\$ en el portafolio 1, al cabo de un año tu retorno será de $3000X_1$\$.

Suponé que las medias y desvíos estandard de X_1, X_2 y X_3 son

	X_1	X_2	X_3
Media	0.08	0.12	0.1
Desvío	0.02	0.05	0.03

Suponé que las correlaciones son

	X_1	X_2	X_3
X_1	1	-0.4	-0.2
X_2	-0.4	1	0.6
X_3	-0.2	0.6	1

(a) Supone que invertís a_j en el portafolio $j = 1, 2, 3$. Como tenés 10000\$, $a_1 + a_2 + a_3 = 10000$. Entonces, tu retorno será $Z = a_1X_1 + a_2X_2 + a_3X_3$.
Encontrá la función $g\,(a_1, a_2, a_3)$ tal que $g\,(a_1, a_2, a_3) = E\,(Z)$

(b) Encontrá la función $h\,(a_1, a_2, a_3)$ tal que $h\,(a_1, a_2, a_3) = Var\,(Z)$

(c) Suponé que sos averso al riesgo, y que querés distribuir tu inversión de modo tal de que el desvío de tu inversión nunca supere el desvío del portafolio con menor desvío, es decir del portafolio 1.
¿Es posible encontrar una estrategia de inversión para la cual el desvío sea menor o igual a $10000 \times 0{,}02 = 200$ (el desvio del retorno si invirtieras 10000\$ en el portafolio 1) pero cuya esperanza de tasa de retorno sea mayor que $10000 \times 0{,}08 = 800$ (la esperanza del retorno si invirtieras todo tu dinero en el portafolio 1)?

(d) Que problema de maximización deberías plantear formalmente para encontrar la mejor estrategia, en el sentido de maximizar la esperanza del retorno sujeto a que el desvío del retorno no supere a 200$?

7. Ejercicio 7

Considere a X una variable aleatoria que puede tomar únicamente los valores -1, 0 y 1 todos con igual probabilidad. Además considere la variable aleatoria Y definida como:

$$Y = \begin{cases} 1 & \text{si } X = 0 \\ 0 & \text{caso contrario.} \end{cases}$$

(a) Calcular la correlación entre X e Y

(b) Son X e Y variables aleatorias independientes?

8. Ejercicio 8

Suponga que Usted desea realizar encuestas sobre las preferencias de los consumidores sobre las tres principales marcas de computadoras portátiles. Para ello usted encuesta a n personas de forma aleatoria, sea la variable X_1 la cantidad de encuestados que eligen la marca 1, la variable X_2 representa la cantidad de personas que eligen la marca 2 y la variable X_3 identifica la cantidad de personas que eligen la marca 3. Se puede pensar entonces que vector (X_1, X_2, X_3) es un vector aleatorio con distribución trinomial, es decir

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{n!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \cdot \mathbb{I}_{\{x_1+x_2+x_3=n\}} \mathbb{I}_{\{x_i \geq 0 \forall i\}}$$

con $p_1 + p_2 + p_3 = 1$. Encuentre la distribución de X_1.

9. Ejercicio 9

La producción de pantallas para computadoras es un proceso difícil que resulta en una proporción importante de pantallas defectuosas. En una compañía, la proporción diaria de pantallas defectuosas de 16 y 20 pulgadas es el resultado de un vector aleatorio bivariado (Z, Y) donde Z es la proporción diaria de pantallas de 16 pulgadas defectuosas e es la proporción diaria de pantallas defectuosas de 20 pulgadas. Supongamos que las proporciones de pantallas defectuosas de la empresa de distintos días están distribuidas de acuerdo a una distribución conjunta que se asume es igual a:

$$f(y, z) = [\theta z + (2 - \theta) y] \mathbb{I}_{(0<z<1)} \mathbb{I}_{(0<y<1)}$$

(a) Encuentre $\mathbb{E}[Z]$.

(b) Encuentre $\mathbb{E}[Y]$.

10. Ejercicio 10

Sean X e Y dos variables aleatorias con una función de densidad conjunta de probabilidad

$$f(x, y) = \frac{2}{3(x + y)} e^{-x} \mathbb{I}_{(0<x)} \mathbb{I}_{(0<y<1)}$$

Encuentre el coeficiente de correlación entre X e Y.

11. Ejercicio 11

La cantidad de personas que entran a una farmacia es una variable aleatoria con distribución Poisson con media λ. De las personas que entran a dicha farmacia la probabilidad de que una mujer entre es p. Encuentre el valor esperado y la varianza de la cantidad de mujeres que entran a la farmacia. ¿Qué supuesto debió realizar? Usando la función generatriz de momentos encuentre la distribución de la cantidad de mujeres que entran a una farmacia.

12. Ejercicio 12

Considere $f(x|y) = \frac{1}{y}e^{-\frac{x}{y}} \cdot \mathbb{I}_{\mathbb{R}_+}(x)$ donde Y es una variable aleatoria con distribución geométrica con probabilidad de éxito es p.

(a) Encuentre $\mathbb{E}[X]$.

(b) Encuentre $\text{Var}(X)$.

13. Ejercicio 13

Sumas aleatorias. Consideremos una suma del estilo

$$T = \sum_{i=1}^{N} X_i$$

donde N es una variable aleatoria con esperanza finita y las X representan variables aleatorias que son independientes entre sí e independientes de N y tienen media común $\mathbb{E}(X)$. Estas sumas aparecen en distintas situaciones:

- Una compañía de seguros podría recibir N reclamos en un periodo de tiempo y los montos de los reclamos individuales pueden modelarse como variables aleatorias $X_1, X_2, \ldots$
- La variable aleatoria N podría representar el número de clientes que ingresan en una tienda y X_i representa los gastos del $i-$ ésimo cliente.
- N podría representar el número de jobs o procesos en la cola de un servidor y X_i podría representar el tiempo de procesamiento del i-ésimo proceso, etc.

Calcule $\mathbb{E}(T)$ y $Var(T)$.

14. Ejercicio 14

Sea $X_i \overset{iid}{\sim} \text{Exp}\left(\frac{1}{\theta}\right)$ para $i = 1, \ldots, n$. Se puede demostrar (lo harán en la siguiente materia) que el estimador de máxima verosimilitud de θ es $\overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$.

(a) Demuestre que $\overline{X}_n$ es un estimador insesgado de θ. Es decir, demuestre que $\mathbb{E}[\overline{X}_n] = \theta$.

(b) Encuentre $\text{Var}(\overline{X}_n)$.

15. Ejercicio 15

Encuentre las distribuciones marginales de X e Y ¿A qué familia de distribuciones pertenecen?

(i) $f(x, y) = [2x + 2y - 4xy]\mathbb{I}_{(0<x<1)}\mathbb{I}_{(0<y<1)}$.

(ii) $f(x,y) = [2-2x-2y+4xy]\mathbb{I}_{(0<x<1)}\mathbb{I}_{(0<y<1)}$.

16. Ejercicio 16

La función de densidad conjunta de X e Y está dada por

$$f(x,y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right)\mathbb{I}_{(0<x<1)}\mathbb{I}_{(0<y<2)}$$

(a) Verifique que de hecho $f(x,y)$ es una función de densidad.

(b) Encuentre la función de densidad de X.

(c) Encuentre $P(X > Y)$.

17. Ejercicio 17

¿Son X e Y variables aleatorias independientes?

(a) $f(x,y) = xe^{-(x+y)}\mathbb{I}_{(0<x)}\mathbb{I}_{(0<y)}$.

(b) $f(x,y) = 2\mathbb{I}_{(0<x<y)}\mathbb{I}_{(0<y<1)}$.

18. Ejercicio 18

Si X tiene distribución uniforme en el intervalo $[0,1]$ e Y es una variable aleatoria con distribución exponencial con media igual a 1. Hallar la distribución de

(a) $Z = X + Y$.

(b) $Z = \frac{X}{Y}$.

19. Ejercicio 19

Sea la función se densidad conjunta de X e Y

$$f(x,y) = e^{-y}\mathbb{I}_{(0<x<y)}$$

(a) Encuentre la covarianza entre X e Y.

(b) Encuentre las funciones de densidad de las variables aleatorias $X|Y$ y $Y|X$. Calcule sus esperanzas.

20. Ejercicio 20

Sean X e Y dos v.a. con distribución exponencial de manera que la CDF conjunta es

$$F_{X,Y}(x,y) = \left(1-e^{-\lambda x}\right)\left(1-e^{-\lambda y}\right), \quad x \geq 0, y \geq 0$$

Encuentre la PDF conjunta y las PDF marginales de cada variable.

21. Ejercicio 21

Supongamos que se arroja una moneda al aire 3 veces. Sea X la v.a. que representa la cantidad de caras que se obtienen en los 3 intentos, mientras que Y es una v.a. que representa la cantidad de caras que se obtienen en los primeros 2 intentos.

Calcule la probabilidad conjunta de (X, Y).

22. Ejercicio 22

Sean X e Y dos variables aleatorias independientes tales que $X, Y \sim \text{Unif}(0,1)$. Encuentre la función de distribución conjunta del vector aleatorio (X, Y).

23. Ejercicio 23

Para la PDF conjunta (vista en la teórica)

$$f(x,y) = \begin{cases} \lambda^2 e^{-\lambda y} & \text{si } 0 \leq x \leq y, \quad \lambda > 0 \\ 0 & \text{en otro caso} \end{cases}$$

muestre que $\mathbb{E}(XY) = \dfrac{3}{\lambda^2}$

Además, encuentre las funciones de densidad marginales de X e Y. Luego, encuentre $f_{X|Y}(x|y)$.

24. Ejercicio 24

Consideremos la siguiente función de densidad bivariada:

$$f(x,y) = \frac{12}{7}\left(x^2 + xy\right)\mathbb{I}_{[0,1]}(x)\mathbb{I}_{[0,1]}(y)$$

Calcule $P(X > Y)$, la función de densidad marginal de X, la función de densidad marginal de Y y las distribuciones condicionales.

25. Ejercicio 25

En este ejercicio no se le pide resolver algo en particular. Es recomendable que lo lea y entienda.

Las carreras ofrecidas en cierta universidad duran cuatro años. Al finalizar cada año, los alumnos rinden un examen general sobre el material de las materias cursadas durante ese año. Las calificaciones posibles, de menor a mayor son $\{1, 2, 3, 4\}$. Elegimos un alumno al azar y registramos X como el año que acaba de finalizar e Y como la nota que obtuvo en el examen final.

$p_{XY}(x,y)$		x (año)				$p_Y(y)$
		1	2	3	4	
	1	0.10	0.05	0.03	0.03	0.21
y	2	0.08	0.10	0.07	0.05	0.30
(nota)	3	0.07	0.06	0.09	0.06	0.28
	4	0.05	0.03	0.05	0.08	0.21
$p_X(x)$		0.30	0.24	0.24	0.22	1

En la siguiente tabla, cada columna registra la función de masa de probabilidad condicional de la nota Y dado cada año x. En la ultima fila, la tabla registra la esperanza de la nota Y dado cada año x

		$p_{Y\|X}(y\|1)$	$p_{Y\|X}(y\|2)$	$p_{Y\|X}(y\|3)$	$p_{Y\|X}(y\|4)$
	1	0.33	0.21	0.125	0.14
y	2	0.27	0.42	0.29	0.23
(nota)	3	0.23	0.25	0.375	0.27
	4	0.17	0.12	0.21	0.36
$E(Y\|X=x)$		**2.24**	**2.28**	**2.67**	**2.85**

donde $E(Y|X=x) = 1 \cdot p_{Y|X}(1|x) + 2 \cdot p_{Y|X}(2|x) + 3 \cdot p_{Y|X}(3|x) + 4 \cdot p_{Y|X}(4|x)$.

La esperanza condicional es útil para resumir la relación que existe entre dos variables Y y X.

x (año)	1	2	3	4
$E(Y\|X=x)$	**2.24**	**2.28**	**2.67**	**2.85**

Vemos que la esperanza de la nota aumenta a medida que avanzan los años de estudio. ¿Por qué?

- Los alumnos mejoran su capacidad de aprendizaje.
- Los alumnos mejoran su concentración para responder en exámenes.
- Los exámenes de los cursos de los años superiores son mas fáciles.
- Los alumnos de los años superiores son los *sobrevivientes* de los años inferiores (notemos que la distribución de X -año- demuestra que hay más alumnos en los años inferiores que superiores)

x (año)	1	2	3	4
$p_X(x)$	0.30	0.24	0.24	0.22

- Los alumnos de los años superiores son mejores porque la calidad de la educación secundaria ha ido empeorando con los años, por lo cual las cohortes superiores vienen mejor preparadas.

Calculemos la esperanza de la nota de un alumno elegido al azar de toda la universidad:

$p_{XY}(x,y)$		x (año)				$p_Y(y)$
		1	2	3	4	
	1	0.10	0.05	0.03	0.03	0.21
y	2	0.08	0.10	0.07	0.05	0.30
(nota)	3	0.07	0.06	0.09	0.06	0.28
	4	0.05	0.03	0.05	0.08	0.21
$p_X(x)$		0.30	0.24	0.24	0.22	1

$$\begin{aligned} E(Y) &= 1 \cdot p_Y(1) + 2 \cdot p_Y(2) + 3 \cdot p_Y(3) + 4 \cdot p_Y(4) \\ &= 1 \cdot 0{,}21 + 2 \cdot 0{,}30 + 3 \cdot 0{,}28 + 4 \cdot 0{,}21 \\ &= 2{,}49 \end{aligned}$$

Notemos que el cálculo de $E(Y)$ se usan los valores de $p_Y(y)$ para todo y. Pero supongamos que no conocemos la función $p_Y(y)$, aunque sí:

1. el número de alumnos en los distintos años
2. el promedio de las notas de todos los alumnos en cada año por separado

Entonces conocemos, para cada $x \in \{1, 2, 3, 4\}$ el valor de

$$n_x = \# \text{ alumnos en el año } x$$

Con lo cual podemos calcular, para cada $x \in \{1, 2, 3, 4\}$ el valor de

$$\frac{n_x}{n_1 + n_2 + n_3 + n_4} = p_X(x)$$

Notemos que esta última igualdad es válida pues

- $p_X(x)$ es la probabilidad de que un alumno sea del año x
- todos los alumnos tienen *la misma probabilidad* de ser elegidos.

Por otro lado, conocemos, para cada $x \in \{1, 2, 3, 4\}$ el valor de

$$\overline{y}_x = \frac{1 \cdot n_{1,x} + 2 \cdot n_{2,x} + 3 \cdot n_{3,x} + 4 \cdot n_{4,x}}{n_x} = 1 \cdot \frac{n_{1,x}}{n_x} + 2 \cdot \frac{n_{2,x}}{n_x} + 3 \cdot \frac{n_{3,x}}{n_x} + 4 \cdot \frac{n_{4,x}}{n_x}$$

donde $n_{i,x} = \#$ de alumnos en el año x con nota i. Pero

$$\frac{n_{i,x}}{n_x} = p_{Y|X}(i|x)$$

ya que

- $p_{Y|X}(i|x)$ es la probabilidad de que un alumno del año x tenga nota i
- todos los alumnos del año x son equiprobables

Entonces, conocemos, para cada valor $x \in \{1, 2, 3, 4\}$, el valor de

$$\mathbb{E}(Y|X = x) = 1 \cdot p_{Y|X}(1|x) + 2 \cdot p_{Y|X}(1|x) + 3 \cdot p_{Y|X}(1|x) + 4 \cdot p_{Y|X}(1|x)$$

Recapitulemos: supongamos que conocemos

1. $p_X(x)$ para cada año $x \in \{1, 2, 3, 4\}$
2. $\mathbb{E}(Y|X = x)$ para cada año $x \in \{1, 2, 3, 4\}$ y

O sea, conocemos la tabla

x	1	2	3	4
$\mathbb{E}(Y\|X = x)$	2.24	2.28	2.67	2.85
$p_X(x)$	0.30	0.24	0.24	0.22

que nos permite calcular $\mathbb{E}(Y)$

$$\begin{aligned}
\mathbb{E}(Y) &= \mathbb{E}(Y|X = 1)p_X(1) + \mathbb{E}(Y|X = 2)p_X(2) \\
&\quad + \mathbb{E}(Y|X = 3)p_X(3) + \mathbb{E}(Y|X = 4)p_X(4) \\
&= 2{,}24 \cdot 0{,}30 + 2{,}28 \cdot 0{,}24 + 2{,}67 \cdot 0{,}24 + 2{,}85 \cdot 0{,}22 \\
&= 2{,}49
\end{aligned}$$

Intuitivamente, la cuenta que estamos haciendo es:

- El promedio global $\overline{y}$ de todas las notas de la universidad es el *promedio ponderado* de los promedios $\overline{y}_x$ de las notas de cada grupo $x = 1, 2, 3, 4$, donde la ponderación es proporcional al tamaño n_x del grupo x.

- Como en nuestro ejemplo Y representa la nota de un alumno elegido al azar de la universidad, y todos los alumnos tienen igual probabilidad de ser elegidos, entonces

$$\begin{aligned}\mathbb{E}(Y) &= \bar{y} \text{ (promedio global)} \\ \mathbb{E}(Y|X=x) &= \bar{y}_x \text{ (promedio del grupo } x) \\ p_X(x) &= n_x/n \text{ (ponderacion del grupo } x)\end{aligned}$$

Varianza condicional

Así como calculamos la esperanza de la nota Y para cada año x, también podemos calcular la varianza de la nota Y para cada año x.

		$p_{Y\|X}(y\|1)$	$p_{Y\|X}(y\|2)$	$p_{Y\|X}(y\|3)$	$p_{Y\|X}(y\|4)$
	1	0.33	0.21	0.125	0.14
y	2	0.27	0.42	0.29	0.23
(nota)	3	0.23	0.25	0.375	0.27
	4	0.17	0.12	0.21	0.36
$E(Y\|X=x)^2$		$2{,}24^2$	$2{,}28^2$	$2{,}67^2$	$2{,}85^2$
$E(Y^2\|X=x)$		6.2	6.06	8.02	9.25
$Var(Y\|X=x)$		**1.18**	**0.86**	**0.89**	**1.13**

Por ejemplo,

$$\mathbb{E}(Y^2|X=1) = 1^2 \cdot 0{,}33 + 2^2 \cdot 0{,}27 + 3^2 \cdot 0{,}23 + 4^2 \cdot 0{,}17 = 6{,}2$$

y

$$Var(Y|X=1) = \mathbb{E}(Y^2|X=1) - [\mathbb{E}(Y|X=1)]^2 = 6{,}2 - 2{,}24^2 = 1{,}1824$$

Independencia

En nuestro ejemplo, hemos calculado

		$p_{Y\|X}(y\|1)$	$p_{Y\|X}(y\|2)$	$p_{Y\|X}(y\|3)$	$p_{Y\|X}(y\|4)$	$p_Y(y)$
	1	0.33	0.21	0.125	0.14	**0.21**
y	2	0.27	0.42	0.29	0.23	**0.30**
(nota)	3	0.23	0.25	0.375	0.27	**0.28**
	4	0.17	0.12	0.21	0.36	**0.21**

Entonces, en nuestro ejemplo, cualquiera sea el x, vimos que

$$p_{Y|X}(y|x) \neq p_Y(y) \text{ para al menos un valor de } y$$

De modo que enterarnos en qué año está el alumno sorteado cambia el pronóstico sobre las posibles notas que se sacó. O sea, la distribución condicional de Y dado $X = x$ no coincide con la distribución de Y.

Pero ahora supongamos que la distribución conjunta de (X, Y) fuera como la que viene dada en la siguiente tabla

		x (año)				$p_Y(y)$
		1	2	3	4	
	1	0.075	0.06	0.06	0.055	**0.25**
y	2	0.1125	0.09	0.09	0.0825	**0.375**
(nota)	3	0.075	0.06	0.06	0.055	**0.25**
	4	0.0375	0.03	0.03	0.0275	**0.125**
$p_X(x)$		0.30	0.24	0.24	0.22	1

En este caso,

$$p_{Y|X}(y|x) = p_Y(y) \quad \forall x, y$$

Por ejemplo,

$$p_{Y|X}(2|4) = \frac{p_{XY}(4,2)}{p_X(4)} = \frac{0{,}03}{0{,}24} = 0{,}125 = p_Y(4)$$

En este ejemplo, la tabla de probabilidades condicionales $p_{Y|X}(y|x)$ y la de probabilidad marginal $p_Y(y)$ es

		$p_{Y\|X}(y\|1)$	$p_{Y\|X}(y\|2)$	$p_{Y\|X}(y\|3)$	$p_{Y\|X}(y\|4)$	$p_Y(y)$
	1	0.25	0.25	0.25	0.25	**0.25**
y	2	0.375	0.375	0.375	0.375	**0.375**
(nota)	3	0.25	0.25	0.25	0.25	**0.25**
	4	0.125	0.125	0.125	0.125	**0.125**

O sea, todas las columnas son iguales. Cuando esto ocurre, decimos que X e Y son variables aleatorias **independientes**. Las siguientes tres afirmaciones son equivalentes en cuanto a independencia de variables aleatorias

1. $p_{Y|X}(y|x) = p_Y(y)$ para todo $x \in Sop(X)$ e $y \in Sop(Y)$
2. $p_{XY}(x,y) = p_X(x)p_Y(y)$ para todo $x \in Sop(\text{X})$ e $y \in Sop(Y)$
3. $p_{X|Y}(x|y) = p_X(x)$ para todo $x \in Sop(X)$ e $y \in Sop(Y)$

con $Sop(X)$ e $Sop(Y)$ los soportes de X e Y respectivamente.