409 TEXTOS DOCENTS

ESTADÍSTICA PARA LAS CIENCIAS DEL COMPORTAMIENTO
Ejercicios comentados

Antonio Solanas
David Leiva
Rumen Manolov
Maribel Peró
Joan Guàrdia

Departamento de Psicología Social
y Psicología Cuantitativa

UNIVERSITAT DE BARCELONA

ESTADÍSTICA PARA LAS CIENCIAS DEL COMPORTAMIENTO

Ejercicios comentados

Antonio Solanas
David Leiva
Rumen Manolov
Maribel Peró
Joan Guàrdia

Departamento de Psicología Social
y Psicología Cuantitativa

409 TEXTOS DOCENTS

UNIVERSITAT DE BARCELONA
Edicions

Índice

Introducción

El término «estadística» se suele asociar con valores numéricos y representaciones gráficas que expresan un resumen de unos hechos o fenómenos de interés. La primera tarea de un lector u oyente crítico no es quedarse fascinado por la sofisticación cuantitativa o visual, sino preguntarse sobre dos aspectos fundamentales: a) cómo se han obtenido estos resúmenes y b) qué significan exactamente. De hecho, las mismas cuestiones se presentan delante de un profesional que ha obtenido datos y necesita organizarlos de alguna manera para que se conviertan en «información». En la época actual, dominada por la tecnología y las aplicaciones informáticas, la obtención de los resúmenes estadísticos parece no acarrear ninguna dificultad, gracias al software estadístico disponible, tanto de pago como gratuito. Sin embargo, la pregunta de cómo se obtienen los resultados está inherentemente relacionada con el tipo de datos de los que se dispone y a los índices que se han utilizado. Nos referimos, en este caso, a uno de los aspectos que vertebran el presente documento: las escalas de medida de las variables. ¿Es suficiente con averiguar que los índices calculados son apropiados para los datos disponibles? Es necesario, pero no es suficiente. La posibilidad de obtener resúmenes numéricos de manera inmediata y sin más esfuerzo que varios clics en el software correspondiente puede ser de gran ayuda, pero también puede representar una dificultad para el segundo de los aspectos mencionados anteriormente: la interpretación de los resultados. El software facilita y agiliza la tarea del analista de datos, pero lo aleja del proceso, lo que convierte el paso de los datos brutos a los resultados en una caja negra con contenido desconocido. En este sentido, consideramos que conocer exactamente las operaciones que se han realizado con los datos puede ayudar a entender qué información dan los resúmenes numéricos (por ejemplo, en el caso de la descripción estadística univariante, si los indicadores son resistentes o no a valores anómalos, si se basan en la totalidad de las medidas disponibles o solo en un porcentaje de los datos centrales de la distribución ordenada).

Partiendo de esta convicción, el objetivo del presente documento es invitar al lector a que sea activo a la hora de resolver las cuestiones planteadas en relación con la estadística descriptiva e inferencial. Se sugiere que la resolución de los problemas se realice de manera autónoma antes de contrastarla con las soluciones ofrecidas, tanto a nivel numérico como a nivel verbal. La extensión de estas soluciones responde a la finalidad de guiar no solo el proceso de cálculo, sino el proceso de reflexión sobre el cálculo y el resultado obtenido. Por tanto, los comentarios que acompañan al cálculo hacen referencia tanto a la obtención de valores numéricos como a su interpretación. En este punto, es preciso dejar claro que la interpretación de los valores proporcionados por los índices puede complementarse y verse favorecida por las representaciones gráficas que resumen los mismos resultados, pero de una forma visual, un aspecto que hemos considerado oportuno resaltar en algunos de los ejercicios. No obstante, el presente documento no se centra en la elaboración e interpretación de gráficos, debido a la amplia disponibilidad de textos sobre la materia, empezando por el texto clásico de John Tukey sobre el análisis exploratorio de datos y, más recientemente, continuando por los libros escritos por Edward Tufte.

La organización del presente documento es la siguiente: en primer lugar, se ofrecen ejercicios sobre conceptos básicos de probabilidad. La finalidad que se persigue es que dichos conceptos estén activados en la mente de los lectores a la hora de enfrentarse a los problemas relacionados con las distribuciones de probabilidad. En este primer bloque, hemos decidido centrarnos en detalle en la distribución binomial y en la ley normal (entre otros modelos de probabilidad), abarcando, por ende, aspectos de variables tanto discretas como continuas. En segundo lugar, se presentan los problemas relacionados con la estadística descriptiva, comenzando por la descripción univariante para posteriormente trabajar diferentes maneras de evaluar y cuantificar la asociación entre dos variables. En este segundo bloque, se hace patente la importancia de la escala de medida de las variables a la hora de escoger los índices estadísticos apropiados. En tercer lugar, una vez tratadas las bases de la probabilidad y la estadística descriptiva, se abordan cuestiones propias de la estadística inferencial. Este último bloque comprende una gran variedad de pruebas estadísticas, tanto paramétricas como no paramétricas. Adicionalmente, se incluyen ejercicios sobre el modelado de relaciones entre variables mediante el análisis de la regresión.

Teniendo en cuenta el contenido del documento, que incluye información práctica sobre el proceso de cálculo, referencias a aspectos teóricos y enfatiza el aprendizaje de interpretación de los resultados estadísticos, consideramos que el presente texto puede resultar de utilidad en cursos de introducción a la estadística. De hecho, el documento puede ser utilizado tanto por quienes se inician en la estadística, permitiéndoles un aprendizaje autónomo, como por los docentes, que podrán aprovechar el material para las sesiones prácticas. Para enriquecer tanto el aprendizaje autónomo como las sesiones presenciales en las que se ilustran los contenidos teóricos, sugerimos trabajar el presente documento conjuntamente con las actividades prácticas incluidas en un recurso digital (Manolov, Solanas, Leiva, Losada, Peró y Guàrdia, 2015), que se centra en los mismos bloques que se tratan aquí.

PARTE 1
VARIABLES ALEATORIAS Y PROBABILIDAD

- ¿Cómo se obtiene conocimiento a partir de la incertidumbre?
- ¿Qué información proporcionan los modelos de probabilidad?
- ¿Qué condición se requiere, necesariamente, para obtener información válida mediante los modelos de probabilidad?
- ¿Qué es lo que distingue a las variables discretas de las continuas?

Estas y muchas otras preguntas encontrarán respuesta a continuación.

Los problemas que se presentan en este primer bloque repasan los conceptos básicos relacionados con la probabilidad (por ejemplo, la unión e intersección de sucesos, la independencia de eventos y la probabilidad condicionada) previamente a adentrarse en los modelos de probabilidad para variables aleatorias discretas y continuas, donde se tratan, entre otros, los siguientes aspectos: la obtención de momentos y sus propiedades, la posibilidad o no de obtener probabilidades para valores concretos de las variables aleatorias, la manera de obtener la probabilidad para uno o más intervalos de valores y la inversa de la función de distribución. De los modelos de probabilidad para variables discretas, se abordan la distribución binomial y la distribución de Poisson, mientras que, para variables continuas, el énfasis se pone en la distribución normal y la distribución uniforme, aunque muchas de las maneras de proceder que se muestran a continuación son aplicables al resto de los modelos discretos y continuos. Algunos de los conceptos expuestos en este primer bloque son fundamentales en la inferencia estadística, como ocurre, por ejemplo, con el valor p, también denominado «grado de significación», asociado al «estadístico de prueba» cuando se pretenden realizar extrapolaciones para las poblaciones de referencia a partir de los resultados obtenidos en muestras.

1. NOCIONES DE PROBABILIDAD

Ejercicio 1

Dado un espacio muestral A definido como $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, justifique cuál de las siguientes alternativas se corresponde con una función de probabilidad que designa probabilidades al conjunto de sucesos elementales.

a) $\Pr(a_1) = 1/6$; $\Pr(a_2) = 1/6$; $\Pr(a_3) = 1/6$; $\Pr(a_4) = 1/6$; $\Pr(a_5) = 1/6$; $\Pr(a_6) = 1/6$

b) $\Pr(a_1) = 1/6$; $\Pr(a_2) = -1/6$; $\Pr(a_3) = 2/6$; $\Pr(a_4) = 1/6$; $\Pr(a_5) = 1/6$; $\Pr(a_6) = 1/6$

c) $\Pr(a_1) = 0{,}3$; $\Pr(a_2) = 0{,}1$; $\Pr(a_3) = 0{,}15$; $\Pr(a_4) = 0{,}25$; $\Pr(a_5) = 0{,}1$; $\Pr(a_6) = 0{,}2$

d) $\Pr(a_1) = 0{,}4$; $\Pr(a_2) = -0{,}1$; $\Pr(a_3) = 0{,}2$; $\Pr(a_4) = 0{,}3$; $\Pr(a_5) = 0{,}1$; $\Pr(a_6) = 0{,}1$

e) $\Pr(a_1) = 2/6$; $\Pr(a_2) = 1/6$; $\Pr(a_3) = 1/6$; $\Pr(a_4) = 2/6$; $\Pr(a_5) = 1/6$; $\Pr(a_6) = 1/6$

Solución. De las expresiones anteriores, la única que se constituye como espacio muestral es la a, puesto que la suma de las probabilidades asignadas a los 6 sucesos elementales, $\Pr(a_i)$, es igual a 1 y todas las probabilidades de los sucesos elementales $\Pr(a_i)$ presentan valores entre 0 y 1.

La expresión b no es un espacio muestral puesto que $\Pr(a_2)$ presenta un valor negativo, aspecto que vulnera el axioma de no negatividad. La expresión c no es un espacio muestral porque, a pesar de que todas las $\Pr(a_i)$ están entre 0 y 1, la suma de todas ellas es superior a 1; en concreto, esta suma devuelve un resultado igual a 1,1 y, por tanto, se vulnera el axioma de la aditividad para sucesos mutuamente excluyentes. La expresión d no es un espacio muestral puesto que $\Pr(a_2)$ presenta un valor negativo, aspecto que vulnera el axioma de la no negatividad. Finalmente, la expresión e tampoco es un espacio de probabilidad puesto que, a pesar de que todas las $\Pr(a_i)$ están entre 0 y 1, la suma de las mismas es superior a la unidad (8/6), aspecto que vulnera el axioma de la normalización.

Ejercicio 2

Si $\Omega = \{a_1, a_2, a_3, a_4\}$ es un espacio muestral, hállese el valor de probabilidad asociado al suceso elemental a_1 para las siguientes expresiones:

a) si $\Pr(a_2) = 0{,}25$, $\Pr(a_3) = 0{,}25$ y $\Pr(a_4) = 0{,}25$

b) si $\Pr(a_2) = 1/4$, $\Pr(a_3) = 1/8$ y $\Pr(a_4) = 1/2$

c) si $\Pr(a_2) = \Pr(a_3) = 0{,}25$ y $\Pr(a_4) = 2 \cdot \Pr(a_1)$

Solución. Teniendo en cuenta que la probabilidad asociada al espacio muestral Ω es igual a la unidad, la suma de las probabilidades asociadas a los cuatro sucesos elementales debe ser igual a la unidad (axioma de normalización). Por consiguiente:

$$\Pr(a_1)+\Pr(a_2)+\Pr(a_3)+\Pr(a_4)=1$$

Así pues, para resolver la expresión a se deben seguir los siguientes pasos:

$$\begin{aligned}\Pr(a_1)+0{,}25+0{,}25+0{,}25 &= 1\\ \Pr(a_1)+0{,}75 &= 1\\ \Pr(a_1)=1-0{,}75 &= 0{,}25.\end{aligned}$$

Los pasos a seguir para resolver la expresión b se resumen a continuación. Puesto que

$$\Pr(a_1)+1/4+1/8+1/2=1$$

y que

$$\frac{1}{4}+\frac{1}{8}+\frac{1}{2}=\frac{2+1+4}{8}=\frac{7}{8}=0{,}875,$$

se puede escribir

$$\begin{aligned}\Pr(a_1)+0{,}875 &= 1\\ \Pr(a_1)=1-0{,}875 &= 0{,}125.\end{aligned}$$

Finalmente, los pasos a seguir para resolver la expresión c se muestran a continuación. Si

$$\Pr(a_2)=\Pr(a_3)=0{,}25 \quad \text{y} \quad \Pr(a_4)=2\cdot\Pr(a_1),$$

entonces

$$\begin{aligned}\Pr(a_1)+0{,}25+0{,}25+2\cdot\Pr(a_1) &= 1\\ 3\cdot\Pr(a_1)+0{,}5 &= 1\\ 3\cdot\Pr(a_1) &= 1-0{,}5\\ 3\cdot\Pr(a_1) &= 0{,}5\\ \Pr(a_1) &= 0{,}5/3\approx 0{,}1667.\end{aligned}$$

Ejercicio 3

En una clase de 35 niños, 10 tienen problemas escolares. Contéstense las siguientes cuestiones.

a) La probabilidad de que, al seleccionar 3 niños al azar, los 3 tengan problemas escolares.

Solución. Puesto que en el aula hay 35 niños y 10 tienen problemas escolares, se puede resolver aplicando la definición clásica de probabilidad. Por tanto:

$$\Pr(x=3)=\frac{\text{casos favorables}}{\text{casos posibles}}$$

La obtención tanto del número de casos favorables como del número de casos posibles se realiza a partir de la utilización de números combinatorios. Así pues, el número de casos favorables se obtiene de la combinación de los 10 niños con problemas escolares que existen en el aula seleccionados de 3 en 3. Así:

$$C_n^k=\binom{n}{k}=\binom{10}{3}=\frac{10!}{3!\cdot 7!}=\frac{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}=120.$$

El número de casos posibles se obtiene de la combinación de los 35 niños del aula seleccionados en grupos de tres. Por tanto:

$$C_n^k=\binom{n}{k}=\binom{35}{3}=\frac{35!}{3!\cdot 32!}=\frac{35\cdot 34\cdot 33\cdot 32\cdot\cdots\cdot 1}{3\cdot 2\cdot 1\cdot 32\cdot 31\cdot 30\cdot 29\cdot\cdots\cdot 1}=6{,}545.$$

Una vez que se dispone del número de casos favorables y del número de casos posibles, obtener la probabilidad de que, al seleccionar 3 niños al azar, los 3 tengan problemas escolares es:

$$\Pr(X=3)=\frac{120}{6.545}\approx 0{,}0183.$$

Otra manera de solucionar esta cuestión es:

$$\Pr(X=3)=\frac{10}{35}\cdot\frac{9}{34}\cdot\frac{8}{33}\approx 0{,}0183.$$

La lógica subyacente a este planteamiento es la siguiente: antes de seleccionar el primer alumno, los casos favorables que cumplen la característica de tener problemas escolares son 10 y existen 35 casos posibles. Al seleccionar el segundo alumno quedan 34 casos posibles, pero, a la vez, de los casos favorables, existen 9, puesto que se supone que el primer niño seleccionado tenía problemas escolares. Lo mismo sucede al seleccionar al tercer niño. De los 33 casos posibles que quedan en el aula, solo 8 pueden ser casos favorables. Puesto que los sucesos son independientes, para obtener la probabilidad de que 3 tengan problemas escolares, se calcula el producto de estas tres probabilidades.

b) La probabilidad de que, al seleccionar 2 niños al azar, ninguno de ellos tenga problemas escolares.

Solución. En este caso se trata de obtener la probabilidad de que los 2 niños seleccionados no tengan problemas escolares. Al igual que en el caso anterior, se puede obtener esta probabilidad a partir de la definición clásica de probabilidad:

$$\Pr(x=k)=\frac{\text{casos favorables}}{\text{casos posibles}}$$

Nuevamente, el número de casos favorables y el de casos posibles se puede obtener a partir de números combinatorios. Así, el número de casos favorables se obtiene a partir de la combinación de 25 niños que no tienen problemas escolares seleccionados de 2 en 2:

$$C_n^k = \binom{n}{k} = \binom{25}{2} = \frac{25!}{2! \cdot 23!} = \frac{25 \cdot 24 \cdot 23 \cdot \cdots \cdot 1}{2 \cdot 1 \cdot 23 \cdot 22 \cdot 21 \cdot \cdots \cdot 1} = 300.$$

El número de casos posibles se determina de la combinación de los 35 niños del aula seleccionados en grupos de 2. Por tanto:

$$C_n^k = \binom{n}{k} = \binom{35}{2} = \frac{35!}{2! \cdot 33!} = \frac{35 \cdot 34 \cdot 33 \cdot \cdots \cdot 1}{2 \cdot 1 \cdot 33 \cdot 32 \cdot 31 \cdot \cdots \cdot 1} = 595.$$

Una vez que se dispone del número de casos favorables y de casos posibles, obtener la probabilidad de que, al seleccionar a 2 niños al azar, los 2 no tengan problemas escolares es:

$$\Pr(X = 2) = \frac{300}{595} \approx 0{,}5042.$$

Otra manera de solucionar esta cuestión es:

$$\Pr(X = 2) = \frac{25}{35} \cdot \frac{24}{34} \approx 0{,}5042.$$

La lógica subyacente a este planteamiento es la siguiente: al seleccionar al primer alumno, los casos favorables que cumplen la característica de no tener problemas escolares son 25 y existen 35 casos posibles. Cuando se va a seleccionar al segundo alumno, quedan 34 casos posibles, pero, a la vez, de los casos favorables quedan 24, puesto que se supone que el primer niño seleccionado era un niño del subgrupo sin problemas escolares. Puesto que los sucesos son independientes, para obtener la probabilidad de que los 2 no tengan problemas escolares se realiza el producto de estas dos probabilidades.

Ejercicio 4

En una muestra de 70 personas, 20 tienen un diagnóstico de solo depresión, 10 tienen un diagnóstico únicamente de ludopatía y 15 tienen un diagnóstico de comorbilidad, es decir, tanto depresión como ludopatía. Teniendo en cuenta esta información, resuélvanse las siguientes cuestiones:

a) El número de personas con un diagnóstico de depresión en la muestra.

Solución. Puesto que 20 personas de la muestra tienen solo diagnóstico de depresión y 15 tienen un diagnóstico tanto de depresión como de ludopatía, el número total de personas que tienen depresión en la muestra estudiada es de 35.

b) El número de personas con un diagnóstico de ludopatía en la muestra.

Solución. Puesto que 10 personas de la muestra tienen solo diagnóstico de ludopatía y 15 tienen un diagnóstico tanto de depresión como de ludopatía, el número total de personas que tienen un diagnóstico de ludopatía en la muestra estudiada es de 25.

c) El número de personas sin diagnóstico en la muestra.

Solución. El número total de personas que tienen algún diagnóstico es de 45 (20 solo depresión $+10$ solo ludopatía $+15$ depresión $\cap$ ludopatía). Por tanto, el número total de personas de la muestra sin diagnóstico es de $70-45$, es decir, 25.

d) La probabilidad de tener un diagnóstico de depresión teniendo en cuenta la información de la muestra.

Solución. Se aplica la definición empírica de probabilidad. Por consiguiente:

$$\Pr(A) = \frac{n_A}{n}$$

En este caso el número total de medidas realizadas es 70 ($n = 70$) y el número total de personas con diagnóstico de depresión es de 35 ($n_A = 35$). Por tanto, aplicando la expresión anterior, la probabilidad de tener un diagnóstico de depresión es de 0,5, tal como se muestra a continuación:

$$\Pr(\text{depresión}) = \frac{35}{70} = 0{,}5.$$

e) La probabilidad de tener un diagnóstico de ludopatía teniendo en cuenta la información de la muestra.

Solución. Se aplica la definición empírica de probabilidad. Así:

$$\Pr(A) = \frac{n_A}{n}$$

En este caso el número total de medidas realizadas es 70 ($n = 70$) y el número total de personas con diagnóstico de ludopatía es de 25 ($n_A = 25$). De este modo, aplicando la expresión anterior, la probabilidad de tener un diagnóstico de depresión es aproximadamente igual a 0,3571, tal como se muestra a continuación:

$$\Pr(\text{ludopatía}) = \frac{25}{70} \approx 0{,}3571.$$

f) La probabilidad de tener un diagnóstico de depresión y de ludopatía teniendo en cuenta la información de la muestra.

Solución. Es necesario encontrar la probabilidad asociada a la intersección de dos sucesos (depresión ∩ ludopatía). Del enunciado del ejercicio se sabe que 15 personas de la muestra cumplen este criterio. Se aplicará la definición empírica de probabilidad:

$$\Pr(A \cap B) = \frac{n_{(A\cap B)}}{n}.$$

En este caso el número total de medidas realizadas es 70 ($n = 70$) y el número total de personas con diagnóstico de depresión y ludopatía es de 15 ($n_{(A\cap B)} = 15$). Así, aplicando la expresión anterior, la probabilidad de tener un diagnóstico de depresión y ludopatía es de casi 0,2143, como se obtiene del siguiente cálculo:

$$\Pr(\text{depresión} \cap \text{ludopatía}) = \frac{15}{70} \approx 0{,}2143.$$

g) La probabilidad de tener un diagnóstico de depresión o de ludopatía teniendo en cuenta la información de la muestra.

Solución. Es necesario encontrar la probabilidad asociada a la unión de dos sucesos (depresión ∪ ludopatía). Por tanto, se tiene que aplicar la expresión:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

De los apartados d), e) y f) se sabe el valor asignado a estas tres probabilidades: $\Pr(\text{depresión}) = 0{,}5$, $\Pr(\text{ludopatía}) \approx 0{,}3571$ y $\Pr(\text{depresión} \cap \text{ludopatía}) \approx 0{,}2143$. Por consiguiente:

$$\Pr(\text{depresión} \cup \text{ludopatía}) \approx 0{,}5 + 0{,}3571 - 0{,}2143 = 0{,}6428.$$

En este caso, como sabemos que el número total de personas que tienen diagnóstico de depresión o ludopatía es de 45, también podríamos resolver el ejercicio aplicando la definición empírica de probabilidad y, por tanto, sería:

$$\Pr(\text{depresión} \cup \text{ludopatía}) = \frac{45}{70} \approx 0{,}6428.$$

h) La probabilidad de tener un diagnóstico de depresión si se sabe que se tiene un diagnóstico de ludopatía.

Solución. En este caso se trabaja con un espacio muestral reducido, puesto que ya se sabe que la persona tiene un diagnóstico de ludopatía. En concreto, se trata de una probabilidad condicionada. En consecuencia, se debe aplicar la siguiente expresión:

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

La probabilidad asociada a la intersección de tener un diagnóstico de depresión y un diagnóstico de ludopatía es 0,2143 (ver apartado f) y la probabilidad de tener un diagnóstico de ludopatía es de 0,3571 (ver apartado e). En consecuencia, la probabilidad de tener un diagnóstico de depresión si se sabe que se tiene un diagnóstico de ludopatía es:

$$\Pr(\text{depresión} \mid \text{ludopatía}) \approx \frac{0{,}2143}{0{,}3571} \approx 0{,}6001.$$

i) La probabilidad de tener un diagnóstico de ludopatía si se sabe que se tiene un diagnóstico de depresión.

Solución. Nuevamente, en este caso se trabaja con un espacio muestral reducido, puesto que ya se sabe que la persona tiene un diagnóstico de depresión. En concreto, se trata de una probabilidad condicionada. En consecuencia, se debe aplicar la siguiente expresión:

$$\Pr(A \mid B) = \frac{Pr(A \cap B)}{\Pr(B)}$$

La probabilidad asociada a la intersección de tener un diagnóstico de depresión y un diagnóstico de ludopatía es 0,2143 (ver apartado f), y la probabilidad de tener un diagnóstico de depresión es de 0,5 (ver apartado d). En consecuencia, la probabilidad de tener un diagnóstico de ludopatía si se sabe que se tiene un diagnóstico de depresión es:

$$\Pr(\text{ludopatía} \mid \text{depresión}) \approx \frac{0{,}2143}{0{,}5} = 0{,}4286.$$

Ejercicio 5

Según diversos estudios realizados en la población española, la probabilidad de no sufrir insomnio o dificultades para dormir es de 0,55 (sin problemas de sueño), la probabilidad de padecer un trastorno de dependencia al alcohol es de 0,099 y la probabilidad de tener problemas de sueño y sufrir un trastorno de dependencia al alcohol es de 0,05.

a) Teniendo en cuenta la información anterior, determinar la probabilidad de no sufrir un trastorno de dependencia al alcohol.

Solución. En este caso se trata de obtener la probabilidad asociada al complementario de sufrir un trastorno de dependencia al alcohol. Por tanto, la expresión para obtener esta probabilidad es la siguiente:

$$\Pr(A)^c = 1 - \Pr(A)$$

Puesto que la probabilidad de sufrir un trastorno de dependencia al alcohol es de 0,099, $\Pr(A)$, su complementario es:

$$\Pr(\text{dependencia alcohol})^c = 1 - 0{,}099 = 0{,}901.$$

b) Determinar la probabilidad de sufrir un problema de sueño o sufrir un problema de dependencia al alcohol.

Solución. En este caso se trata de obtener la probabilidad asociada a la unión de dos sucesos no disjuntos. Por tanto, la expresión de cálculo es la siguiente:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

En el enunciado se proporciona la probabilidad de no sufrir un problema de sueño (0,55), pero, puesto que se pide la de la unión de sufrir un problema de sueño y padecer un problema de dependencia al alcohol, primero se debe obtener la probabilidad de sufrir un problema de sueño, que se puede calcular aplicando la ley del complementario:

$$\Pr(\bar{A}) \equiv \Pr(A)^c = 1 - \Pr(A)$$

En la expresión anterior, $\bar{A}$ significa que no se da el suceso A, que, para el caso que nos ocupa, es que no sufre un problema de sueño. Así pues, la probabilidad de sufrir un problema de sueño es:

$$\Pr(\text{problema sueño}) = 1 - \Pr(\text{problema sueño})^c = 1 - 0{,}55 = 0{,}45.$$

Por consiguiente, ahora ya se dispone de todas las probabilidades necesarias para poder determinar la probabilidad asociada a la unión de los dos sucesos: Pr(problema sueño) = 0,45, Pr(dependencia alcohol) = 0,099 y Pr(problema sueño ∩ dependencia alcohol) = 0,05. Así:

$$\Pr(\text{problema sueño} \cup \text{dependencia alcohol}) = 0{,}45 + 0{,}099 - 0{,}05 = 0{,}499.$$

c) Determinar la probabilidad del complementario de sufrir un problema de sueño y sufrir un problema de dependencia al alcohol.

Solución. En este caso se trata de obtener la probabilidad asociada al complementario de la intersección de dos sucesos. Por tanto, la expresión de cálculo para obtener esta probabilidad es:

$$\Pr(A \cap B)^c = 1 - \Pr(A \cap B)$$

La probabilidad asociada a la intersección de sufrir un problema de sueño y sufrir un problema de dependencia al alcohol es de 0,05. En consecuencia, su complementario (i. e., la probabilidad de no sufrir problemas de sueño ni de dependencia al alcohol al mismo tiempo) es:

$$\Pr(\text{trastorno sueño} \cap \text{dependencia alcohol})^c = 1 - 0{,}05 = 0{,}95.$$