Ejercicios Inferencia Unidad 3. SOLUCIONES

Alex Sanchez y Santi Perez-Hoyos

2024-11-08

Table of contents

1 Ejercicio 1 **2**
1.1 Parte a: Función de probabilidad conjunta para Y_1 y Y_2 . . . 2
1.2 Parte b: Cálculo de $F(1,0)$ . . . 3
1.3 Resumen . . . 3

2 Ejercicio 2 **3**
2.1 1. Función de probabilidad conjunta para Y_1 y Y_2 . . . 4
2.1.1 Probabilidad de cada combinación de (Y_1, Y_2): . . . 4
2.2 2. Cálculo de $F(2,1)$ . . . 4
2.3 Resumen de resultados . . . 5

3 Ejercicio 3 **5**
3.1 Parte a: Distribución de probabilidad marginal de Y_1 . . . 5
3.2 Parte b: Comparación con la distribución binomial . . . 6
3.3 Conclusión . . . 6

4 Ejercicio 4 **7**
4.1 1. Encontrar el valor de k que haga de ésta una función de densidad de probabilidad 7
4.1.1 Paso 1: Determinar el área de la región triangular . . . 7
4.1.2 Paso 2: Integrar $f(y_1, y_2)$ sobre la región triangular . . . 8
4.2 2. Encontrar $P(Y_1 \geq 3Y_2)$ . . . 8
4.2.1 Paso 1: Identificar la subregión definida por $Y_1 \geq 3Y_2$ . . . 8
4.2.2 Paso 2: Calcular el área de la subregión . . . 8
4.2.3 Paso 3: Calcular la probabilidad . . . 8
4.3 Respuesta final . . . 9

5 Ejercicio 5 **9**
5.1 1. Probabilidad de que la cantidad de contaminante en la chimenea con limpiador sea mayor que 0.5 . . . 9
5.1.1 Paso 1: Establecer los límites de integración . . . 9
5.1.2 Paso 2: Integrar la función de densidad conjunta . . . 10
5.1.3 Resultado . . . 10
5.2 2. Probabilidad condicional dada una observación de $Y_2 = 0.5$ . . . 10
5.2.1 Paso 1: Identificar la función de densidad condicional . . . 10
5.2.2 Paso 2: Calcular la probabilidad condicional . . . 11
5.3 Respuesta final . . . 11

6 Ejercicio 6 **11**
6.1 Parte a: Cálculo de la covarianza $\text{Cov}(Y_1, Y_2)$ 12
6.2 Parte b: Interpretación de la covarianza negativa 13
6.3 Ejercicio 13
6.3.1 1. Probabilidad de que la cantidad de contaminante en la chimenea con limpiador sea mayor que 0.5 13
6.3.2 2. Probabilidad condicional dada una observación de $Y_2 = 0.5$ 14
6.3.3 Respuesta final 15

7 Ejercicio 7 **15**
7.1 Parte a: Cálculo de $\text{Cov}(Y_1, Y_1)$ 16
7.2 Parte b: Verificación de $\text{Cov}(Y_1, Y_2) = 7$ 16
7.3 Parte c: Valor máximo posible de $\text{Cov}(Y_1, Y_2)$ y su interpretación 16

8 Ejercicio 8 **17**
8.1 Parte a: Probabilidad de obtener $Y_1 = 3$, $Y_2 = 1$ y $Y_3 = 2$ en $n = 6$ corridas 17
8.2 Parte b: Esperanza y varianza de Y_1 para un n general 17
8.3 Parte c: Covarianza entre Y_2 y Y_3 para un n general 18
8.4 Parte d: Esperanza y varianza de $Y_2 - Y_3$ para un n general 18
8.5 Resumen de Resultados 18

1 Ejercicio 1

Se tienen dos estudios clínicos importantes, cuyos análisis genéticos deben ser asignados aleatoriamente a uno o más de tres laboratorios, A, B y C . Denote con Y_1 el número de estudios asignados al laboratorio A y con Y_2 el número de estudios asignados al laboratorio B. Cada laboratorio puede recibir 0,1 o 2 estudios. a. Encuentre la función de probabilidad conjunta para Y_1 y Y_2. b. Encuentre $F(1, 0)$, es decir, la probabilidad de que el laboratorio A reciba como máximo un estudio y el laboratorio B no reciba ninguno.

1.1 Parte a: Función de probabilidad conjunta para Y_1 y Y_2

En este ejercicio, se nos indica que existen tres laboratorios (A, B y C) a los cuales se pueden asignar los estudios de forma aleatoria. Denotamos con Y_1 el número de estudios asignados al laboratorio A y con Y_2 el número de estudios asignados al laboratorio B. Cada laboratorio puede recibir entre 0 y 2 estudios.

Vamos a analizar el espacio muestral, S, que representa las posibles combinaciones de asignación de estudios a los laboratorios. Los resultados posibles son:

S	AA	AB	AC	BA	BB	BC	CA	CB	CC
(y_1, y_2)	(2,0)	(1,1)	(1,0)	(1,1)	(0,2)	(1,0)	(1,0)	(0,1)	(0,0)

Cada punto muestral es igualmente probable, con una probabilidad de $\frac{1}{9}$, ya que existen 9 combinaciones posibles.

La función de probabilidad conjunta para Y_1 y Y_2 queda entonces representada en la siguiente tabla:

	$y_1 = 0$	$y_1 = 1$	$y_1 = 2$
$y_2 = 0$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$
$y_2 = 1$	$\frac{2}{9}$	$\frac{2}{9}$	0
$y_2 = 2$	$\frac{1}{9}$	0	0

1.2 Parte b: Cálculo de $F(1, 0)$

Nos piden encontrar la probabilidad de que el laboratorio A reciba como máximo un estudio y el laboratorio B no reciba ninguno, es decir, $F(1, 0) = P(Y_1 \leq 1, Y_2 = 0)$.

Para resolverlo, sumamos las probabilidades de los eventos en los cuales $Y_1 \leq 1$ y $Y_2 = 0$, que son $(Y_1 = 0, Y_2 = 0)$ y $(Y_1 = 1, Y_2 = 0)$:

$$F(1, 0) = P(Y_1 = 0, Y_2 = 0) + P(Y_1 = 1, Y_2 = 0)$$

Sustituyendo con las probabilidades correspondientes de la tabla obtenemos:

$$F(1, 0) = \frac{1}{9} + \frac{2}{9} = \frac{3}{9} = \frac{1}{3}$$

1.3 Resumen

1. La función de probabilidad conjunta ha sido obtenida en función de todas las combinaciones posibles de asignación, considerando que cada una es igualmente probable.
2. La probabilidad solicitada, $F(1, 0)$, es de $\frac{1}{3}$, que representa la probabilidad de que el laboratorio A reciba como máximo un estudio y el laboratorio B no reciba ninguno.

2 Ejercicio 2

Tres monedas balanceadas se lanzan en forma independiente al aire. Una de las variables de interés es Y_1, el número de caras. Denote con Y_2 la cantidad de dinero ganado en una apuesta colateral en la siguiente forma. Si la primera cara aparece en el primer tiro, usted gana 1€. Si la primera cara aparece en el tiro segundo o en el tercero gana 2€ o 3€, respectivamente. Si no aparece una cara, usted pierde 1€ (esto es, gana - 1€).

1. Encuentre la función de probabilidad conjunta para Y_1 y Y_2.
2. ¿Cuál es la probabilidad de que haya menos de tres caras y usted gane 1€ o menos? [Esto es, encuentre $F(2, 1)$.

Solución

2.1 1. Función de probabilidad conjunta para Y_1 y Y_2

Dado que se lanzan tres monedas balanceadas de manera independiente, cada lanzamiento puede resultar en cara (C) o cruz (+) con probabilidad 0.5.

Listamos todas las posibles secuencias de resultados en los tres lanzamientos y calculamos los valores de Y_1 (número de caras obtenidas) y Y_2 (cantidad de dinero ganado) para cada caso.

Secuencia	Y_1 (Número de Caras)	Y_2 (Ganancia en €)
CCC	3	1
CC+	2	1
C+C	2	1
C++	1	1
+CC	2	2
+C+	1	2
++C	1	3
+++	0	-1

Para calcular la función de probabilidad conjunta $P(Y_1 = y_1, Y_2 = y_2)$, obtenemos las probabilidades de cada combinación de (Y_1, Y_2) a partir de la cantidad de secuencias que cumplen con esos valores específicos.

2.1.1 Probabilidad de cada combinación de (Y_1, Y_2):

- $P(Y_1 = 3, Y_2 = 1) = \frac{1}{8}$, secuencia: CCC
- $P(Y_1 = 2, Y_2 = 1) = \frac{2}{8}$, secuencias: CC+, C+C
- $P(Y_1 = 2, Y_2 = 2) = \frac{1}{8}$, secuencia: +CC
- $P(Y_1 = 1, Y_2 = 1) = \frac{1}{8}$, secuencia: C++
- $P(Y_1 = 1, Y_2 = 2) = \frac{1}{8}$, secuencia: +C+
- $P(Y_1 = 1, Y_2 = 3) = \frac{1}{8}$, secuencia: ++C
- $P(Y_1 = 0, Y_2 = -1) = \frac{1}{8}$, secuencia: +++

Con esto, la función de probabilidad conjunta $P(Y_1 = y_1, Y_2 = y_2)$ se define por:

$$P(Y_1 = y_1, Y_2 = y_2) = \begin{cases} \frac{1}{8}, & (y_1, y_2) = (3,1), (2,1), (2,2), (1,1), (1,2), (1,3), (0,-1) \\ 0, & \text{en cualquier otro caso} \end{cases}$$

2.2 2. Cálculo de $F(2,1)$

Buscamos $F(2,1) = P(Y_1 \leq 2, Y_2 \leq 1)$, es decir, la probabilidad de obtener menos de tres caras y ganar un máximo de 1€.

Para calcular esta probabilidad, sumamos $P(Y_1 = y_1, Y_2 = y_2)$ para todos los pares (y_1, y_2) que cumplen $Y_1 \leq 2$ y $Y_2 \leq 1$. De la tabla anterior, vemos que cumplen esta condición las siguientes combinaciones:

- $(Y_1 = 2, Y_2 = 1)$ con probabilidad $\frac{2}{8}$
- $(Y_1 = 1, Y_2 = 1)$ con probabilidad $\frac{1}{8}$

- $(Y_1 = 0, Y_2 = -1)$ con probabilidad $\frac{1}{8}$

Entonces:

$$F(2,1) = P(Y_1 \leq 2, Y_2 \leq 1) = \frac{2}{8} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

2.3 Resumen de resultados

1. La función de probabilidad conjunta $P(Y_1 = y_1, Y_2 = y_2)$ se ha especificado para todos los valores posibles.
2. La probabilidad de que haya menos de tres caras y se gane 1€ o menos es $F(2,1) = \frac{1}{2}$.

3 Ejercicio 3

En el Ejercicio 1 determinamos que la distribución conjunta de Y_1, el número de análisis asignados al laboratorio A , y Y_2, el número de análisis asignados al laboratorio B , está dada por las entradas en la siguiente tabla.

y_1			
y_2	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

a. Encuentre la distribución de probabilidad marginal de Y_1.
b. De acuerdo con los resultados vistos anteriormente Y_1 tiene una distribución binomial con $n = 2$ y $p = 1/3$. ¿Hay algún conflicto entre este resultado y la respuesta dada en el punto a?

Solución

3.1 Parte a: Distribución de probabilidad marginal de Y_1

Para encontrar la distribución marginal de Y_1, debemos sumar las probabilidades conjuntas de Y_1 y Y_2 para cada valor posible de Y_1. A continuación, mostramos la tabla de probabilidades conjuntas de la solución anterior:

$y_1 \backslash y_2$	0	1	2
$y_1 = 0$	1/9	2/9	1/9
$y_1 = 1$	2/9	2/9	0
$y_1 = 2$	1/9	0	0

La distribución marginal de Y_1 se calcula sumando las probabilidades de cada fila (fijando y_1 y sumando sobre los valores de y_2):

1. Para $y_1 = 0$:
$$P(Y_1 = 0) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

2. Para $y_1 = 1$:
$$P(Y_1 = 1) = \frac{2}{9} + \frac{2}{9} + 0 = \frac{4}{9}$$

3. Para $y_1 = 2$:
$$P(Y_1 = 2) = \frac{1}{9} + 0 + 0 = \frac{1}{9}$$

Por lo tanto, la distribución marginal de Y_1 es:

y_1	0	1	2
$p_{Y_1}(y_1)$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

3.2 Parte b: Comparación con la distribución binomial

Se sugiere que Y_1 sigue una distribución binomial con parámetros $n = 2$ y $p = \frac{1}{3}$. Veamos si esta afirmación concuerda con los resultados obtenidos en la parte a.

Para una variable aleatoria binomial $Y_1 \sim \text{Binomial}(n = 2, p = 1/3)$, la función de probabilidad es:

$$P(Y_1 = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Calculamos las probabilidades para cada valor de k:

1. Para $k = 0$:
$$P(Y_1 = 0) = \binom{2}{0} \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^2 = 1 \cdot 1 \cdot \frac{4}{9} = \frac{4}{9}$$

2. Para $k = 1$:
$$P(Y_1 = 1) = \binom{2}{1} \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^1 = 2 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{4}{9}$$

3. Para $k = 2$:
$$P(Y_1 = 2) = \binom{2}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^0 = 1 \cdot \frac{1}{9} \cdot 1 = \frac{1}{9}$$

Las probabilidades obtenidas para la binomial son exactamente las mismas que encontramos en la distribución marginal de Y_1, lo que confirma que $Y_1 \sim \text{Binomial}(2, 1/3)$.

3.3 Conclusión

No hay conflicto entre la distribución marginal de Y_1 obtenida en la parte a y el hecho de que Y_1 tenga una distribución binomial con parámetros $n = 2$ y $p = \frac{1}{3}$.

4 Ejercicio 4

Un ingeniero ambiental mide la cantidad (en peso) de partículas contaminantes en muestras de aire de cierto volumen recolectado en dos chimeneas en una planta de energía alimentada con carbón. Una de las chimeneas está equipada con un aparato limpiador. Denote con Y_1 la cantidad de contaminante por muestra recolectada arriba de la chimenea que no tiene aparato limpiador y denote con Y_2 la cantidad de contaminante por muestra recolectada arriba de la chimenea que está equipada con el aparato limpiador.

Suponga que el comportamiento de frecuencia relativa de Y_1 y Y_2 puede ser modelado por

$$f(y_1, y_2) = \begin{cases} k, & 0 \leq y_1 \leq 2, \quad 0 \leq y_2 \leq 1, \quad 2y_2 \leq y_1 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Esto es, Y_1 y Y_2 están uniformemente distribuidas sobre la región dentro del triángulo limitado por $y_1 = 2, y_2 = 0$ y $2y_2 = y_1$.

1. Encuentre el valor de k que haga de ésta una función de densidad de probabilidad.
2. Encuentre $P(Y_1 \geq 3Y_2)$. Esto es, encuentre la probabilidad de que el aparato limpiador reduzca la cantidad de contaminante en un tercio o más.

Solución

4.1 1. Encontrar el valor de k que haga de ésta una función de densidad de probabilidad

Para que $f(y_1, y_2)$ sea una función de densidad de probabilidad válida, la integral de $f(y_1, y_2)$ sobre toda la región de soporte debe ser igual a 1:

$$\iint_{\text{región de soporte}} f(y_1, y_2)\, dy_1\, dy_2 = 1$$

La función de densidad es constante y toma el valor k sobre la región triangular definida por $0 \leq y_1 \leq 2$, $0 \leq y_2 \leq 1$ y $2y_2 \leq y_1$. Esta región corresponde a un triángulo en el plano y_1-y_2 con los vértices en $(0, 0)$, $(2, 0)$, y $(2, 1)$.

4.1.1 Paso 1: Determinar el área de la región triangular

La región triangular tiene una base de longitud 2 (a lo largo del eje y_1) y una altura de 1 (a lo largo del eje y_2). El área del triángulo es:

$$\text{Área} = \frac{1}{2} \times \text{base} \times \text{altura} = \frac{1}{2} \times 2 \times 1 = 1$$

4.1.2 Paso 2: Integrar $f(y_1, y_2)$ sobre la región triangular

Dado que $f(y_1, y_2) = k$ en esta región y la función de densidad es uniforme, la integral sobre esta área es simplemente k multiplicado por el área:

$$\iint_{\text{región}} f(y_1, y_2)\, dy_1\, dy_2 = k \times \text{Área} = k \times 1 = k$$

Para que $f(y_1, y_2)$ sea una función de densidad, necesitamos que esta integral sea igual a 1, entonces:

$$k = 1$$

4.2 2. Encontrar $P(Y_1 \geq 3Y_2)$

Queremos encontrar la probabilidad de que $Y_1 \geq 3Y_2$ en la región triangular donde $f(y_1, y_2) = k = 1$.

4.2.1 Paso 1: Identificar la subregión definida por $Y_1 \geq 3Y_2$

La desigualdad $Y_1 \geq 3Y_2$ corresponde a la recta $y_1 = 3y_2$. Para encontrar la intersección de esta recta con la región triangular, notamos que: - En $y_1 = 2$, al sustituir en $y_1 = 3y_2$, tenemos $y_2 = \frac{2}{3}$.

Por lo tanto, la subregión de interés es el triángulo delimitado por los puntos $(0, 0)$, $(2, 0)$ y $(2, \frac{2}{3})$.

4.2.2 Paso 2: Calcular el área de la subregión

La base de este subtriángulo es 2 (a lo largo de y_1), y la altura es $\frac{2}{3}$ (a lo largo de y_2). Su área es:

$$\text{Área del subtriángulo} = \frac{1}{2} \times \text{base} \times \text{altura} = \frac{1}{2} \times 2 \times \frac{2}{3} = \frac{2}{3}$$

4.2.3 Paso 3: Calcular la probabilidad

La probabilidad buscada es la proporción del área del subtriángulo respecto al área total de la región de soporte:

$$P(Y_1 \geq 3Y_2) = \frac{\text{Área del subtriángulo}}{\text{Área de la región total}} = \frac{\frac{2}{3}}{1} = \frac{2}{3}$$

4.3 Respuesta final

1. El valor de k que hace de $f(y_1, y_2)$ una función de densidad de probabilidad es $k = 1$.
2. La probabilidad de que $Y_1 \geq 3Y_2$ es:

$$P(Y_1 \geq 3Y_2) = \frac{2}{3}$$

5 Ejercicio 5

En el Ejercicio 4 hemos establecido que

$$f\left(y_1, y_2\right) = \begin{cases} k, & 0 \leq y_1 \leq 2, \quad 0 \leq y_2 \leq 1, \quad 2y_2 \leq y_1 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

es una función de densidad de probabilidad conjunta válida para Y_1, la cantidad de contaminante por muestra recolectada arriba de la chimenea que no tenía el aparato limpiador, y para Y_2, la cantidad recolectada arriba de la chimenea con el aparato limpiador.

1. Si consideramos la chimenea con el limpiador instalado, encuentre la probabilidad de que la cantidad de contaminante en una muestra determinada sea mayor que 0.5.

b. Dado que se observa que la cantidad de contaminante en una muestra tomada arriba de la chimenea con el limpiador es 0.5 , encuentre la probabilidad de que la cantidad de contaminante exceda de 1.5 arriba de la otra chimenea (la que no tiene limpiador).

Solución

5.1 1. Probabilidad de que la cantidad de contaminante en la chimenea con limpiador sea mayor que 0.5

Queremos encontrar la probabilidad de que $Y_2 > 0.5$, donde Y_2 representa la cantidad de contaminante en la chimenea con el aparato limpiador. Recordando que $f(y_1, y_2)$ es la función de densidad conjunta y que hemos hallado $k = 1$, calcularemos la probabilidad integrando sobre la región correspondiente a $y_2 > 0.5$ y y_1 dentro de su rango definido por la condición $2y_2 \leq y_1 \leq 2$.

La probabilidad se calcula como:

$$P(Y_2 > 0.5) = \iint_{\{y_2 > 0.5, 2y_2 \leq y_1 \leq 2\}} f(y_1, y_2)\, dy_1\, dy_2$$

5.1.1 Paso 1: Establecer los límites de integración

Para $y_2 > 0.5$, los límites de y_1 están restringidos por la región triangular dada: - $2y_2 \leq y_1 \leq 2$

Entonces, los límites de integración son: - y_2: desde 0.5 hasta 1 - y_1: desde $2y_2$ hasta 2

5.1.2 Paso 2: Integrar la función de densidad conjunta

Dado que $f(y_1, y_2) = 1$ en esta región, la probabilidad es la integral de 1 sobre el área triangular correspondiente:

$$P(Y_2 > 0.5) = \int_{0.5}^{1} \int_{2y_2}^{2} 1 \, dy_1 \, dy_2$$

Evaluamos esta integral en dos pasos.

```r
# Calculo de la probabilidad con R

# Definir la función de integración para y1
integrate_y1 <- function(y2) {
  integrate(function(y1) 1, lower = 2 * y2, upper = 2)$value
}

# Integrar respecto a y2
result <- integrate(function(y2) integrate_y1(y2), lower = 0.5, upper = 1)
result$value
```

5.1.3 Resultado

Al resolver esta integral, obtenemos la probabilidad:

$$P(Y_2 > 0.5) \approx 0.25$$

5.2 2. Probabilidad condicional dada una observación de $Y_2 = 0.5$

Queremos calcular $P(Y_1 > 1.5 \mid Y_2 = 0.5)$, la probabilidad de que la cantidad de contaminante en la chimenea sin limpiador (representada por Y_1) sea mayor que 1.5, dado que en la chimenea con limpiador se observó $Y_2 = 0.5$.

5.2.1 Paso 1: Identificar la función de densidad condicional

Para la función de densidad condicional $f_{Y_1|Y_2}(y_1|y_2)$, aplicamos:

$$f_{Y_1|Y_2}(y_1|Y_2 = 0.5) = \frac{f(y_1, 0.5)}{f_{Y_2}(0.5)}$$

Calcularemos $f_{Y_2}(0.5)$ integrando $f(y_1, y_2)$ sobre los valores de y_1 en la región donde $Y_2 = 0.5$:

$$f_{Y_2}(0.5) = \int_{y_1=2\cdot 0.5}^{2} f(y_1, 0.5) \, dy_1 = \int_{1}^{2} 1 \, dy_1$$

Evaluamos la integral:

$$f_{Y_2}(0.5) = \int_1^2 1\, dy_1 = (2-1) = 1$$

Por lo tanto, $f_{Y_2}(0.5) = 1$.

5.2.2 Paso 2: Calcular la probabilidad condicional

La probabilidad de interés es:

$$P(Y_1 > 1.5 | Y_2 = 0.5) = \int_{1.5}^2 f_{Y_1|Y_2}(y_1|Y_2 = 0.5)\, dy_1$$

Como $f_{Y_1|Y_2}(y_1|Y_2 = 0.5) = 1$ para $1 \leq y_1 \leq 2$, tenemos:

$$P(Y_1 > 1.5 | Y_2 = 0.5) = \int_{1.5}^2 1\, dy_1 = 2 - 1.5 = 0.5$$

5.3 Respuesta final

1. La probabilidad de que la cantidad de contaminante en la chimenea con limpiador sea mayor que 0.5 es aproximadamente 0.25.
2. La probabilidad de que la cantidad de contaminante en la chimenea sin limpiador exceda 1.5, dado que en la otra chimenea se observó 0.5, es 0.5.

6 Ejercicio 6

En el ejercicio 1 determinamos que la distribución conjunta de Y_1, el número de análisis asignados al laboratorio A, y Y_2, el número de análisis asignados al laboratorio B , está dada por las entradas en la siguiente tabla.

	y_1		
y_2	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

a. Encuentre $\mathrm{Cov}\,(Y_1, Y_2)$. $_i$
b. Le sorprende que $\mathrm{Cov}\,(Y_1, Y_2)$ sea negativa? ¿Por qué?

Solución

6.1 Parte a: Cálculo de la covarianza $\text{Cov}(Y_1, Y_2)$

La covarianza entre dos variables aleatorias Y_1 y Y_2 se define como:

$$\text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2)$$

Para calcular la covarianza, necesitamos obtener los valores de $E(Y_1)$, $E(Y_2)$ y $E(Y_1 Y_2)$.

1. **Cálculo de $E(Y_1)$ y $E(Y_2)$**

   A partir de la distribución conjunta dada en la tabla, podemos calcular la esperanza de Y_1 y Y_2 sumando las posibles combinaciones de valores ponderadas por sus probabilidades.

$$E(Y_1) = \sum_{y_1} y_1 \cdot P(Y_1 = y_1)$$
$$E(Y_2) = \sum_{y_2} y_2 \cdot P(Y_2 = y_2)$$

   Usamos la tabla para estos cálculos:

```
# Probabilidades conjuntas
probs <- matrix(c(1/9, 2/9, 1/9, 2/9, 2/9, 0, 1/9, 0, 0), nrow = 3, byrow = TRUE)

# Valores de Y1 y Y2
y1_values <- 0:2
y2_values <- 0:2

# Esperanza de Y1
E_Y1 <- sum(y1_values * rowSums(probs))
# Esperanza de Y2
E_Y2 <- sum(y2_values * colSums(probs))

E_Y1
```

```
[1] 0.6666667
```

```
E_Y2
```

```
[1] 0.6666667
```

2. **Cálculo de $E(Y_1 Y_2)$**

   Para calcular $E(Y_1 Y_2)$, sumamos el producto $y_1 \cdot y_2$ ponderado por la probabilidad conjunta $p(y_1, y_2)$.

```
# Producto de Y1 * Y2 * probabilidad conjunta
E_Y1Y2 <- sum(outer(y1_values, y2_values, "*") * probs)

E_Y1Y2
```

```
[1] 0.2222222
```

3. **Calcular la covarianza**

   Sustituyendo los valores obtenidos:

$$\text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2)$$

```
Cov_Y1Y2 <- E_Y1Y2 - E_Y1 * E_Y2
Cov_Y1Y2
```

```
[1] -0.2222222
```

6.2 Parte b: Interpretación de la covarianza negativa

La covarianza calculada es negativa. Este resultado tiene sentido en el contexto del problema. Dado que los estudios se asignan a tres laboratorios y cada laboratorio recibe un número limitado de estudios, un incremento en el número de estudios asignados a un laboratorio reduce la cantidad disponible para los otros. Así, si Y_1 aumenta, Y_2 tiende a disminuir, lo que explica una relación inversa entre ambas variables y resulta en una covarianza negativa.

Solución

6.3 Ejercicio

6.3.1 1. Probabilidad de que la cantidad de contaminante en la chimenea con limpiador sea mayor que 0.5

Queremos encontrar la probabilidad de que $Y_2 > 0.5$, donde Y_2 representa la cantidad de contaminante en la chimenea con el aparato limpiador. Recordando que $f(y_1, y_2)$ es la función de densidad conjunta y que hemos hallado $k = 1$, calcularemos la probabilidad integrando sobre la región correspondiente a $y_2 > 0.5$ y y_1 dentro de su rango definido por la condición $2y_2 \leq y_1 \leq 2$.

La probabilidad se calcula como:

$$P(Y_2 > 0.5) = \iint_{\{y_2 > 0.5, 2y_2 \leq y_1 \leq 2\}} f(y_1, y_2)\, dy_1\, dy_2$$

6.3.1.1 Paso 1: Establecer los límites de integración

Para $y_2 > 0.5$, los límites de y_1 están restringidos por la región triangular dada: - $2y_2 \leq y_1 \leq 2$

Entonces, los límites de integración son: - y_2: desde 0.5 hasta 1 - y_1: desde $2y_2$ hasta 2

6.3.1.2 Paso 2: Integrar la función de densidad conjunta

Dado que $f(y_1, y_2) = 1$ en esta región, la probabilidad es la integral de 1 sobre el área triangular correspondiente:

$$P(Y_2 > 0.5) = \int_{0.5}^{1} \int_{2y_2}^{2} 1 \, dy_1 \, dy_2$$

Evaluamos esta integral en dos pasos.

```r
# Calculo de la probabilidad con R

# Definir la función de integración para y1
integrate_y1 <- function(y2) {
  integrate(function(y1) 1, lower = 2 * y2, upper = 2)$value
}

# Integrar respecto a y2
result <- integrate(function(y2) integrate_y1(y2), lower = 0.5, upper = 1)
result$value
```

6.3.1.3 Resultado

Al resolver esta integral, obtenemos la probabilidad:

$$P(Y_2 > 0.5) \approx 0.25$$

6.3.2 2. Probabilidad condicional dada una observación de $Y_2 = 0.5$

Queremos calcular $P(Y_1 > 1.5 \mid Y_2 = 0.5)$, la probabilidad de que la cantidad de contaminante en la chimenea sin limpiador (representada por Y_1) sea mayor que 1.5, dado que en la chimenea con limpiador se observó $Y_2 = 0.5$.

6.3.2.1 Paso 1: Identificar la función de densidad condicional

Para la función de densidad condicional $f_{Y_1|Y_2}(y_1|y_2)$, aplicamos:

$$f_{Y_1|Y_2}(y_1|Y_2 = 0.5) = \frac{f(y_1, 0.5)}{f_{Y_2}(0.5)}$$

Calcularemos $f_{Y_2}(0.5)$ integrando $f(y_1, y_2)$ sobre los valores de y_1 en la región donde $Y_2 = 0.5$:

$$f_{Y_2}(0.5) = \int_{y_1=2\cdot 0.5}^{2} f(y_1, 0.5) \, dy_1 = \int_{1}^{2} 1 \, dy_1$$

Evaluamos la integral:

$$f_{Y_2}(0.5) = \int_1^2 1\, dy_1 = (2-1) = 1$$

Por lo tanto, $f_{Y_2}(0.5) = 1$.

6.3.2.2 Paso 2: Calcular la probabilidad condicional

La probabilidad de interés es:

$$P(Y_1 > 1.5 | Y_2 = 0.5) = \int_{1.5}^2 f_{Y_1|Y_2}(y_1|Y_2 = 0.5)\, dy_1$$

Como $f_{Y_1|Y_2}(y_1|Y_2 = 0.5) = 1$ para $1 \leq y_1 \leq 2$, tenemos:

$$P(Y_1 > 1.5 | Y_2 = 0.5) = \int_{1.5}^2 1\, dy_1 = 2 - 1.5 = 0.5$$

6.3.3 Respuesta final

1. La probabilidad de que la cantidad de contaminante en la chimenea con limpiador sea mayor que 0.5 es aproximadamente 0.25.
2. La probabilidad de que la cantidad de contaminante en la chimenea sin limpiador exceda 1.5, dado que en la otra chimenea se observó 0.5, es 0.5.

7 Ejercicio 7

Las variables aleatorias Y_1 y Y_2 son tales que $E(Y_1) = 4, E(Y_2) = -1, V(Y_1) = 2$ y $V(Y_2) = 8$.

1. ¿Cuál es $\text{Cov}(Y_1, Y_1)$?
2. Suponiendo que las medias y las varianzas sean correctas, ¿es posible que $\text{Cov}(Y_1, Y_2) = 7$? [Sugerencia: si $\text{Cov}(Y_1, Y_2) = 7$, ¿cuál es el valor de ρ, el coeficiente de correlación?]
3. Suponiendo que las medias y las varianzas sean correctas, ¿cuál es el máximo valor posible para $\text{Cov}(Y_1, Y_2)$? Si $\text{Cov}(Y_1, Y_2)$ alcanza este valor máximo, ¿qué implica eso acerca de la relación entre Y_1 y Y_2 ?

Solución

7.1 Parte a: Cálculo de $\mathrm{Cov}(Y_1, Y_1)$

Para cualquier variable aleatoria Y, la covarianza de Y consigo misma es igual a su varianza. Es decir:

$$\mathrm{Cov}(Y_1, Y_1) = V(Y_1)$$

Dado que $V(Y_1) = 2$, tenemos que:

$$\mathrm{Cov}(Y_1, Y_1) = 2$$

7.2 Parte b: Verificación de $\mathrm{Cov}(Y_1, Y_2) = 7$

Si se supone que $\mathrm{Cov}(Y_1, Y_2) = 7$, podemos calcular el coeficiente de correlación ρ, que se define como:

$$\rho = \frac{\mathrm{Cov}(Y_1, Y_2)}{\sqrt{V(Y_1) \cdot V(Y_2)}}$$

Sustituyendo los valores dados:

$$\rho = \frac{7}{\sqrt{2 \cdot 8}} = \frac{7}{4} = 1.75$$

Dado que el coeficiente de correlación ρ debe estar en el rango de $-1 \leq \rho \leq 1$, obtener $\rho = 1.75$ es imposible. Esto implica que no es posible que $\mathrm{Cov}(Y_1, Y_2) = 7$ con los valores de varianza y media proporcionados.

7.3 Parte c: Valor máximo posible de $\mathrm{Cov}(Y_1, Y_2)$ y su interpretación

Para determinar el valor máximo posible de $\mathrm{Cov}(Y_1, Y_2)$, consideramos que el valor absoluto del coeficiente de correlación ρ puede ser como máximo 1. Esto ocurre en los casos de correlación lineal perfecta (positiva o negativa). Entonces, el valor máximo de $\mathrm{Cov}(Y_1, Y_2)$ es:

$$\mathrm{Cov}(Y_1, Y_2) = \rho \cdot \sqrt{V(Y_1) \cdot V(Y_2)}$$

Si $\rho = 1$, lo cual indica una asociación lineal positiva perfecta, obtenemos:

$$\mathrm{Cov}(Y_1, Y_2) = 1 \cdot \sqrt{2 \cdot 8} = 4$$

Esto significa que el valor máximo posible de $\mathrm{Cov}(Y_1, Y_2)$ es 4. De manera similar, si $\rho = -1$, el valor mínimo posible de $\mathrm{Cov}(Y_1, Y_2)$ sería -4, indicando una asociación lineal negativa perfecta.

Cuando $\mathrm{Cov}(Y_1, Y_2) = 4$, significa que Y_1 y Y_2 están perfectamente correlacionadas en forma positiva, es decir, existe una relación lineal exacta en la que un aumento en Y_1 siempre corresponde a un aumento proporcional en Y_2.

8 Ejercicio 8

Un experimento de aprendizaje requiere que una rata corra por un laberinto (una red de pasillos) hasta que localice una de tres posibles salidas. La salida 1 presenta una recompensa de alimento, no así las salidas 2 y 3. (Si la rata finalmente selecciona la salida 1 casi siempre, puede tener lugar el aprendizaje.) Denote con Y_i el número de veces que la salida i es seleccionada en corridas sucesivas. Para lo siguiente, suponga que la rata escoge una salida aleatoriamente en cada corrida.

1. Encuentre la probabilidad de que $n = 6$ corridas resulte en $Y_1 = 3, Y_2 = 1$ y $Y_3 = 2$.
2. Para n general, encuentre $E(Y_1)$ y $V(Y_1)$.
3. Encuentre $\text{Cov}(Y_2, Y_3)$ para n general.
4. Para comprobar la preferencia de la rata entre las salidas 2 y 3 , podemos buscar en $Y_2 - Y_3$. Encuentre $E(Y_2 - Y_3)$ y $V(Y_2 - Y_3)$ para n general.

Solución

8.1 Parte a: Probabilidad de obtener $Y_1 = 3$, $Y_2 = 1$ y $Y_3 = 2$ en $n = 6$ corridas

Este problema puede resolverse utilizando la **distribución multinomial**, ya que describe el número de veces que ocurre cada posible resultado en un número fijo de ensayos independientes. Aquí, cada una de las tres salidas tiene la misma probabilidad de ser seleccionada en cada corrida, es decir, $p_1 = p_2 = p_3 = \frac{1}{3}$, y el número total de corridas es $n = 6$. Por lo tanto, estamos interesados en calcular:

$$P(Y_1 = 3, Y_2 = 1, Y_3 = 2) = \frac{6!}{3!\,1!\,2!} \left(\frac{1}{3}\right)^6$$

Calculando esta expresión:

```
n <- 6
p <- 1 / 3
prob <- factorial(n) / (factorial(3) * factorial(1) * factorial(2)) * p^n
prob
```

```
[1] 0.08230453
```

El resultado de esta probabilidad es aproximadamente 0.0823.

8.2 Parte b: Esperanza y varianza de Y_1 para un n general

Para una variable aleatoria multinomial, la esperanza y la varianza de cada conteo se puede calcular usando las propiedades de esta distribución. En particular, para Y_1, tenemos que:

- La **esperanza** es $E(Y_1) = n \cdot p_1$.
- La **varianza** es $V(Y_1) = n \cdot p_1 \cdot (1 - p_1)$.

Como $p_1 = \frac{1}{3}$, obtenemos:

$$E(Y_1) = \frac{n}{3}, \quad V(Y_1) = n \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{2n}{9}$$

8.3 Parte c: Covarianza entre Y_2 y Y_3 para un n general

La **covarianza** entre dos variables en una distribución multinomial, Y_2 y Y_3, se puede calcular como:

$$\text{Cov}(Y_2, Y_3) = -n \cdot p_2 \cdot p_3$$

Dado que $p_2 = p_3 = \frac{1}{3}$, tenemos:

$$\text{Cov}(Y_2, Y_3) = -n \cdot \frac{1}{3} \cdot \frac{1}{3} = -\frac{n}{9}$$

8.4 Parte d: Esperanza y varianza de $Y_2 - Y_3$ para un n general

Para evaluar la preferencia de la rata entre las salidas 2 y 3, podemos observar la diferencia $Y_2 - Y_3$. Calcularemos su esperanza y varianza.

1. **Esperanza de $Y_2 - Y_3$**:

   Usamos la linealidad de la esperanza:

   $$E(Y_2 - Y_3) = E(Y_2) - E(Y_3) = \frac{n}{3} - \frac{n}{3} = 0$$

2. **Varianza de $Y_2 - Y_3$**:

   Para calcular la varianza de $Y_2 - Y_3$, usamos la fórmula de la varianza de una diferencia:

   $$V(Y_2 - Y_3) = V(Y_2) + V(Y_3) - 2\,\text{Cov}(Y_2, Y_3)$$

   Sabemos que $V(Y_2) = V(Y_3) = \frac{2n}{9}$ y que $\text{Cov}(Y_2, Y_3) = -\frac{n}{9}$, por lo que:

   $$V(Y_2 - Y_3) = \frac{2n}{9} + \frac{2n}{9} - 2 \cdot \left(-\frac{n}{9}\right) = \frac{6n}{9} = \frac{2n}{3}$$

8.5 Resumen de Resultados

1. La probabilidad de observar $Y_1 = 3$, $Y_2 = 1$ y $Y_3 = 2$ en 6 corridas es 0.0823.
2. La esperanza y varianza de Y_1 para n corridas son $E(Y_1) = \frac{n}{3}$ y $V(Y_1) = \frac{2n}{9}$.
3. La covarianza entre Y_2 y Y_3 es $\text{Cov}(Y_2, Y_3) = -\frac{n}{9}$.
4. La esperanza y varianza de $Y_2 - Y_3$ son $E(Y_2 - Y_3) = 0$ y $V(Y_2 - Y_3) = \frac{2n}{3}$.