Probabilidad multivariable

Antonio Falcó

Seminario 6

Empleamos las observaciones obtenidas de una muestra Ω_1 de la talla y el peso de 45 individuos de una población Ω :

	altura	peso
1	180	70
2	177	57
3	180	60
4	180	66
5	183	62
6	184	68
7	185	65
8	184	72
9	174	65
10	180	72
11	168	52
12	180	75
13	183	75
14	181	68
15	180	65

	altura	peso
16	190	66
17	183	78
18	167	60
19	181	67
20	179	98
21	173	75
22	170	68
23	170	59
24	183	72
25	179	73
26	180	72
27	188	70
28	176	65
29	178	72
30	185	71

	altura	peso
31	168	52
32	157	47
33	167	53
34	168	57
35	163	65
36	167	60
37	166	68
38	164	49
39	172	57
40	165	59
41	158	62
42	161	65
43	160	61
44	162	58
45	165	58

Ejemplos

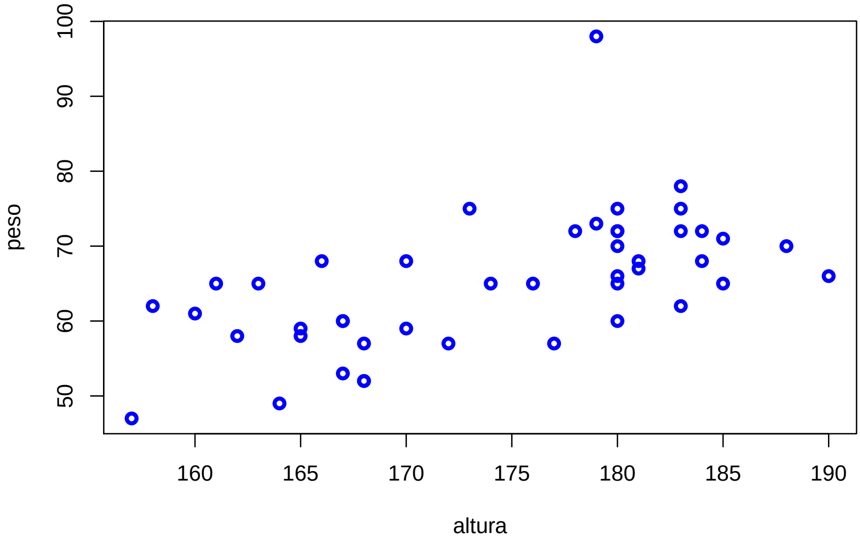
- $\mathbb{E}(\texttt{altura}) = \texttt{mean(altura)} = 174.3111111$ cm,
- $\mathrm{Var}(\texttt{altura}) = \texttt{var(altura)} = 78.0828283$ cm^2,
- $\sqrt{\mathrm{Var}(\texttt{altura})} = \texttt{sd(altura)} = 8.8364489$ cm,
- $\mathbb{E}(\texttt{peso}) = \texttt{mean(peso)} = 65.0888889$ kg,
- $\mathrm{Var}(\texttt{peso}) = \texttt{var(peso)} = 80.1737374$, kg^2.
- $\sqrt{\mathrm{Var}(\texttt{peso})} = \texttt{sd(peso)} = 8.9539789$ kg,

Cuestión científica

Las variables altura y peso están definidas sobre la misma población experimental, podemos entonces representar las dos variables de forma conjunta para cada individuo de la poblacion:

$$\omega \mapsto (\texttt{altura}(\omega), \texttt{peso}(\omega)).$$

¿Podemos afirmar que a mayor altura los individuos de la población tienen un mayor peso?


peso
altura
50
60
70
80
90
100
160
165
170
175
180
185
190

Definición (probabilidad conjunta)

Si en una población experimental $(\Omega, \mathbb{P})$ tenemos una serie de medidas de observación (variables aleatorias) $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ entonces se llama función de densidad conjunta a la función multivariable

$$f_{\mathbf{X}}(y_1, y_2, \ldots, y_n) = \mathbb{P}(\{X_1 = y_1\} \cap \{X_2 = y_2\} \cap \cdots \cap \{X_n = y_n\}),$$

que representa la probabilidad del conjunto de individuos de la población experimental Ω susceptibles de tomar los valores de $X_1 = y_1$, $X_2 = y_2, \ldots, X_n = y_n$. En consecuencia, la función de distribución de $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ se escribe como

$$F(x_1, x_2, \ldots, x_n) = \sum_{y_1 \leq x_1} \sum_{y_2 \leq x_2} \cdots \sum_{y_n \leq x_n} f_{\mathbf{X}}(y_1, y_2, \ldots, y_n).$$

Además se cumple:

$$\sum_{y_1 \in X_1(\Omega)} \sum_{y_2 \in X_2(\Omega)} \cdots \sum_{y_n \in X_n(\Omega)} f_{\mathbf{X}}(y_1, y_2, \ldots, y_n) = 1.$$

Ejemplo

Consideremos $X_1 = \texttt{altura}$ y $X_2 = \texttt{peso}$ definidas sobre una muestra Ω_n de 45-individuos de una población experimental Ω. Entonces $\mathbf{X} = (\texttt{altura}, \texttt{peso})$ y la función de densidad (sobre la muestra) de las dos variables de forma conjunta es:

$$\begin{aligned} f_{\mathbf{X}}(y_1, y_2|\Omega_n) &= \mathbb{P}(\{\texttt{altura} = y_1\} \cap \{\texttt{peso} = y_2\}|\Omega_n) \\ &= \mathbb{P}(\texttt{altura} = y_1, \texttt{peso} = y_2|\Omega_n) \end{aligned}$$

Esta probabilidad en la práctica se puede mostrar en forma de una tabla (matriz).

```
##       peso
## altura 47 49 52 53 57 58 59 60 61 62 65 66 67 68 70 71 72 73 75 78 98
##    157  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    158  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0
##    160  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0
##    161  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0
##    162  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    163  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0
##    164  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    165  0  0  0  0  0  1  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    166  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0
##    167  0  0  0  1  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0
##    168  0  0  2  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    170  0  0  0  0  0  0  1  0  0  0  0  0  0  1  0  0  0  0  0  0  0
##    172  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    173  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0
##    174  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0
##    176  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0
##    177  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    178  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0
##    179  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  1
##    180  0  0  0  0  0  0  0  1  0  0  1  1  0  0  1  0  2  0  1  0  0
##    181  0  0  0  0  0  0  0  0  0  0  0  0  1  1  0  0  0  0  0  0  0
##    183  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  1  0  1  1  0
##    184  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  1  0  0  0  0
##    185  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  1  0  0  0  0  0
##    188  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0
```

```
##        peso
## altura    47    49    52    53    57    58    59    60    61    62    65
##    157 0.022 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    158 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022 0.000
##    160 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022 0.000 0.000
##    161 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022
##    162 0.000 0.000 0.000 0.000 0.000 0.022 0.000 0.000 0.000 0.000 0.000
##    163 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022
##    164 0.000 0.022 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    165 0.000 0.000 0.000 0.000 0.000 0.022 0.022 0.000 0.000 0.000 0.000
##    166 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    167 0.000 0.000 0.000 0.022 0.000 0.000 0.000 0.044 0.000 0.000 0.000
##    168 0.000 0.000 0.044 0.000 0.022 0.000 0.000 0.000 0.000 0.000 0.000
##    170 0.000 0.000 0.000 0.000 0.000 0.000 0.022 0.000 0.000 0.000 0.000
##    172 0.000 0.000 0.000 0.000 0.022 0.000 0.000 0.000 0.000 0.000 0.000
##    173 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    174 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022
##    176 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022
##    177 0.000 0.000 0.000 0.000 0.022 0.000 0.000 0.000 0.000 0.000 0.000
##    178 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    179 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    180 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022 0.000 0.000 0.022
##    181 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    183 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022 0.000
##    184 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##    185 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.022
##    188 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
```

Cuestión

En el seminario 5 construimos las funciones de densidad de las variables individuales $f_{\texttt{altura}}$ y $f_{\texttt{peso}}$. ¿Qué relación existe entre estas probabilidades individuales y la probabilidad conjunta $f_{(\texttt{altura},\texttt{peso})}$?

Respuesta

Conocemos que

$$\sum_{y_1 \in \texttt{altura}(\Omega)} \sum_{y_2 \in \texttt{peso}(\Omega)} f_{(\texttt{altura},\texttt{peso})}(y_1, y_2) = 1,$$

entonces

$$\sum_{y_2 \in \texttt{peso}(\Omega)} f_{(\texttt{altura},\texttt{peso})}(y_1, y_2) = f_{\texttt{altura}}(y_1),$$

y

$$\sum_{y_1 \in \texttt{altura}(\Omega)} f_{(\texttt{altura},\texttt{peso})}(y_1, y_2) = f_{\texttt{peso}}(y_2).$$

Jerga estadística

En estadística a la (multi)-variable $(\texttt{altura}, \texttt{peso})$ se le conoce como **distribución conjunta** y a las variables individuales altura y peso se las conoce como **distribuciones marginales**. En particular, $f_{(\texttt{altura},\texttt{peso})}$ es la función de densidad **conjunta** y $f_{\texttt{altura}}, f_{\texttt{peso}}$ son las funciones de densidad **marginales**.

```
##        peso
## altura 47 49 52 53 57 58 59 60 61 62 65 66 67 68 70 71 72 73 75 78 98 sum
##    157  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0   1
##    158  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0   1
##    160  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0   1
##    161  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0   1
##    162  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0   1
##    163  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0   1
##    164  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0   1
##    165  0  0  0  0  0  1  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0   2
##    166  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0   1
##    167  0  0  0  1  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0   3
##    168  0  0  2  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0   3
##    170  0  0  0  0  0  0  1  0  0  0  0  0  0  1  0  0  0  0  0  0  0   2
##    172  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0   1
##    173  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0   1
##    174  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0   1
##    176  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0   1
##    177  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0   1
##    178  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0   1
##    179  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  1   2
##    180  0  0  0  0  0  0  0  1  0  0  1  1  0  0  1  0  2  0  1  0  0   7
##    181  0  0  0  0  0  0  0  0  0  0  0  0  1  1  0  0  0  0  0  0  0   2
##    183  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  1  0  1  1  0   4
##    184  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  1  0  0  0  0   2
##    185  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  1  0  0  0  0  0   2
##    188  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0   1
```

Definición (variables aleatorias independientes)

Supongamos que sobre una población experimental $(\Omega, \mathbb{P})$ tenemos definida una encuesta $\mathbf{X} = (X_1, \ldots, X_n)$. Diremos que las variables $X_1, X_2, \ldots, X_n$ son independientes si la función de densidad conjunta es el producto de las distribuciones marginales:

$$f_{\mathbf{X}}(y_1, y_2, \ldots, y_n) = f_{X_1}(y_1) f_{X_2}(y_2) \cdots f_{X_n}(y_n).$$

esta expresión es equivalente a

$$\mathbb{P}(\{X_1 = y_1\} \cap \{X_2 = y_2\} \cap \cdots \cap \{X_n = y_n\}) = \mathbb{P}(X_1 = y_1)\mathbb{P}(X_2 = y_2) \cdots \mathbb{P}(X_n = y_n).$$

Ejemplo

Las variables `altura` y `peso` son independientes si

$$f_{(\mathtt{altura},\mathtt{peso})}(y_1, y_2) = f_{\mathtt{altura}}(y_1) f_{\mathtt{peso}}(y_2)$$

se cumple para todas las alturas observadas y_1 y todos los pesos observados y_2.

Independencia

Si la variables altura y peso son independientes entonces se cumple

$$\mathbb{P}(\texttt{altura} = y_1|\texttt{peso} = y_2) = \mathbb{P}(\texttt{altura} = y_1) = f_{\texttt{altura}}(y_1)$$

y

$$\mathbb{P}(\texttt{peso} = y_2|\texttt{altura} = y_1) = \mathbb{P}(\texttt{peso} = y_2) = f_{\texttt{peso}}(y_2)$$

Dependencia

Si la variables altura y peso son dependientes entonces podemos introducir las **funciones de densidad condicionales (o con información a priori):**

$$\mathbb{P}(\texttt{altura} = y_1|\texttt{peso} = y_2) = f_{\texttt{altura}}(y_1|\texttt{peso} = y_2),$$

y

$$\mathbb{P}(\texttt{peso} = y_2|\texttt{altura} = y_1) = f_{\texttt{peso}}(y_2|\texttt{altura} = y_1).$$

Cuestión práctica

¿Cómo podemos determinar si las variables `altura` y `peso` son dependientes o independientes?

Respuesta

Conocemos la distribución de la altura y del peso

	altura	Freq
1	157	0.02
2	158	0.02
3	160	0.02
4	161	0.02
5	162	0.02
6	163	0.02
7	164	0.02
8	165	0.04
9	166	0.02
10	167	0.07
11	168	0.07
12	170	0.04
13	172	0.02
14	173	0.02
15	174	0.02

	altura	Freq
16	176	0.02
17	177	0.02
18	178	0.02
19	179	0.04
20	180	0.16
21	181	0.04
22	183	0.09
23	184	0.04
24	185	0.04
25	188	0.02
26	190	0.02

	peso	Freq
1	47	0.02
2	49	0.02
3	52	0.04
4	53	0.02
5	57	0.07
6	58	0.04
7	59	0.04
8	60	0.07
9	61	0.02
10	62	0.04
11	65	0.13
12	66	0.04
13	67	0.02
14	68	0.09
15	70	0.04

	peso	Freq
16	71	0.02
17	72	0.11
18	73	0.02
19	75	0.07
20	78	0.02
21	98	0.02

Tenemos dos datos, las frecuencias relativas de la variable `altura`

```
zx$Freq
```

```
##  [1] 0.022 0.022 0.022 0.022 0.022 0.022 0.022 0.044 0.022 0.067 0.067
## [12] 0.044 0.022 0.022 0.022 0.022 0.022 0.022 0.044 0.156 0.044 0.089
## [23] 0.044 0.044 0.022 0.022
```

y las frecuencias relativas de la variable `peso`

```
zy$Freq
```

```
##  [1] 0.022 0.022 0.044 0.022 0.067 0.044 0.044 0.067 0.022 0.044 0.133
## [12] 0.044 0.022 0.089 0.044 0.022 0.111 0.022 0.067 0.022 0.022
```

Si fuesen independientes, entonces la tabla de la probabilidad conjunta debería ser

```
##            [,1]    [,2]    [,3]    [,4]   [,5]    [,6]    [,7]   [,8]
##  [1,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [2,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [3,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [4,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [5,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [6,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [7,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
##  [8,] 0.00099 0.00099 0.00198 0.00099 0.0030 0.00198 0.00198 0.0030
##  [9,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [10,] 0.00148 0.00148 0.00296 0.00148 0.0044 0.00296 0.00296 0.0044
## [11,] 0.00148 0.00148 0.00296 0.00148 0.0044 0.00296 0.00296 0.0044
## [12,] 0.00099 0.00099 0.00198 0.00099 0.0030 0.00198 0.00198 0.0030
## [13,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [14,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [15,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [16,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [17,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [18,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
## [19,] 0.00099 0.00099 0.00198 0.00099 0.0030 0.00198 0.00198 0.0030
## [20,] 0.00346 0.00346 0.00691 0.00346 0.0104 0.00691 0.00691 0.0104
## [21,] 0.00099 0.00099 0.00198 0.00099 0.0030 0.00198 0.00198 0.0030
## [22,] 0.00198 0.00198 0.00395 0.00198 0.0059 0.00395 0.00395 0.0059
## [23,] 0.00099 0.00099 0.00198 0.00099 0.0030 0.00198 0.00198 0.0030
## [24,] 0.00099 0.00099 0.00198 0.00099 0.0030 0.00198 0.00198 0.0030
## [25,] 0.00049 0.00049 0.00099 0.00049 0.0015 0.00099 0.00099 0.0015
```

En particular obtenemos que si fuesen independientes se tendría que cumplir:

$$\begin{aligned}\mathbb{P}(\texttt{altura} = 157, \texttt{peso} = 47) &= \mathbb{P}(\texttt{altura} = 157)\mathbb{P}(\texttt{peso} = 47) \\ &= 0.02 \times 0.02 \\ &= 4.94 \times 10^{-4}\end{aligned}$$

Ahora bien, tenemos que

$$\mathbb{P}(\texttt{altura} = 157, \texttt{peso} = 47) = \frac{1}{45} = 0.02 \neq 4.94 \times 10^{-4},$$

en consecuencia podemos afirmar que la variables `altura` y `peso` **NO** son independientes.

Dependencia entre variables

Si dos variables X y Y definidas sobre la misma población experimental $(\Omega, \mathbb{P})$ **NO** son independientes entonces las funciones de densidad condicionales

$$f_X(x|Y=y) = \mathbb{P}(X=x|Y=y) \neq f_X(x) = \mathbb{P}(X=x)$$
$$f_Y(y|X=x) = \mathbb{P}(Y=y|X=x) \neq f_Y(y) = \mathbb{P}(Y=y).$$

son diferentes de las distribuciones marginales.

Valores esperados condicionales

Recordemos que los valores esperados se calculan

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x f_X(x) \quad \mathbb{E}(Y) = \sum_{y \in Y(\Omega)} y f_Y(y).$$

Entonces, podemos introducir los **valores esperados condicionales**:

$$\mathbb{E}(X|Y=y) = \sum_{x \in X(\Omega)} x f_X(x|Y=y) \quad \mathbb{E}(Y|X=x) = \sum_{y \in Y(\Omega)} y f_Y(y|X=x).$$

Ejemplo

Como la variable `altura` y `peso` **NO** son independientes y conocemos que

- $\mathbb{E}(\texttt{altura}) = \texttt{mean(altura)} = 174.31$ cm,
- $\mathrm{Var}(\texttt{altura}) = \texttt{var(altura)} = 78.08$ cm^2,
- $\sqrt{\mathrm{Var}(\texttt{altura})} = \texttt{sd(altura)} = 8.84$ cm,
- $\mathbb{E}(\texttt{peso}) = \texttt{mean(peso)} = 65.09$ kg,
- $\mathrm{Var}(\texttt{peso}) = \texttt{var(peso)} = 80.17$, kg^2.
- $\sqrt{\mathrm{Var}(\texttt{peso})} = \texttt{sd(peso)} = 8.95$ kg,

vamos a calcular sus valores esperados condicionales

Esperanzas condicionales

- $\mathbb{E}(\texttt{altura}|\texttt{peso} = 65) = 173.17$ cm
- $\mathbb{E}(\texttt{altura}|\texttt{peso} = 72) = 181$ cm
- $\mathbb{E}(\texttt{peso}|\texttt{altura} = 180) = 68.57$ kg
- $\mathbb{E}(\texttt{peso}|\texttt{peso} = 168) = 53.67$ kg

Varianzas condicionales

- $\mathrm{Var}(\texttt{altura}|\texttt{peso} = 65) = 89.37$ cm^2
- $\mathrm{Var}(\texttt{altura}|\texttt{peso} = 72) = 6$ cm^2
- $\mathrm{Var}(\texttt{peso}|\texttt{altura} = 180) = 26.62$ kg^2
- $\mathrm{Var}(\texttt{peso}|\texttt{peso} = 168) = 8.33$ kg^2

Relación entre dos variables NO independientes

Si X e Y no son independientes entonces conocemos que

$$\mathbb{E}(Y|X=x)=g(x)$$

es una función que depende de los valores observados x, y del mismo modo

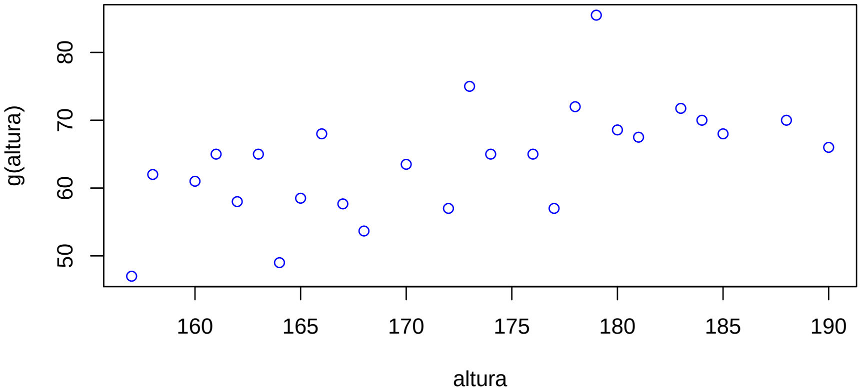
$$\mathbb{E}(X|Y=y)=h(y)$$

es una función que depende de los valores observados y.

Ejemplo

Consideremos las variables $X = \texttt{altura}$ y $Y = \texttt{peso}$, entonces veamos la función del peso esperado en función de la altura observada:

$$g(x) = \mathbb{E}(\texttt{peso}|\texttt{altura} = x)$$



Podemos interpolar entre los puntos: