

Cuestiones Seminarios 6 y 7

Bioestadística

06/05/2020

Cuestiones a responder en el cuaderno de campo

Observaciones y medición - Distribuciones de probabilidad

Consideremos la tabla de datos siguientes que contiene el peso y la altura de una muestra de 45 individuos de una población

	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

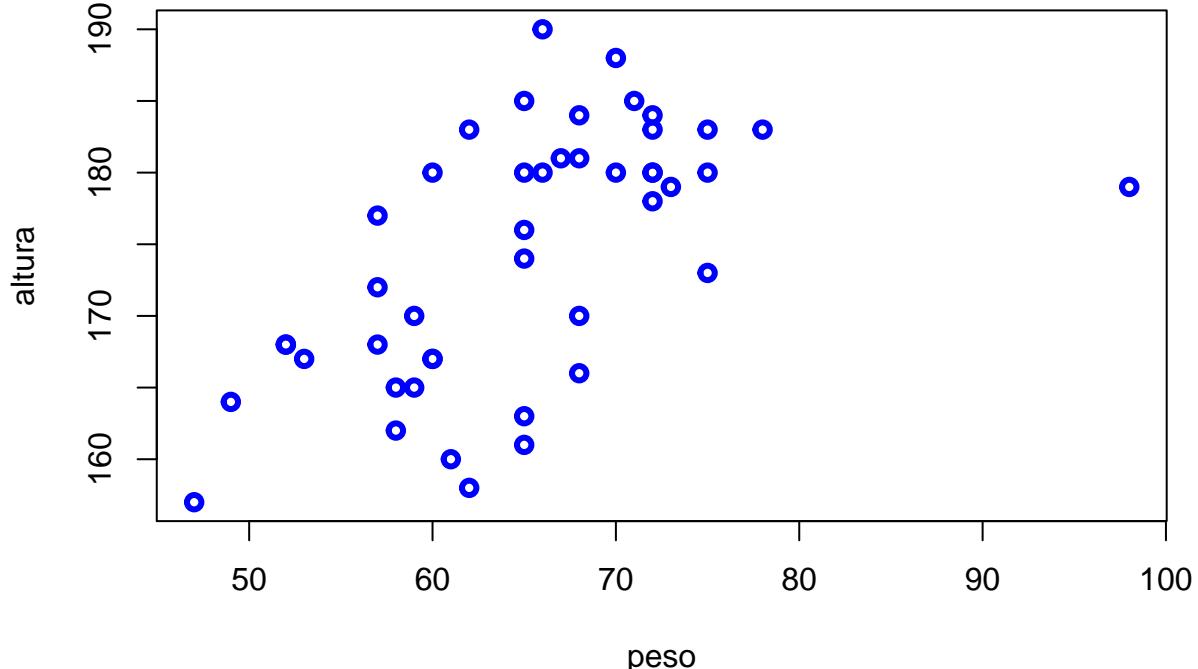
	altura	peso
31	168	52
32	157	47
33	167	53
34	168	57
35	163	65
36	167	60
37	166	68
38	164	49
39	172	57
40	165	59
41	158	62
42	161	65
43	160	61
44	162	58
45	165	58

Cuestión a responder en el cuaderno de campo:

En el seminario hemos estudiado la variable **peso** en relación a la variable **altura**. Vamos a estudiar ahora la relación de la variable **altura** en función de la variable **peso**. La primera cuestión a responder en el cuaderno de campo es la siguiente:

- ¿Cómo podemos justificar el estudio de la relación de la variable **altura** en función de la variable **peso**?

Observar la gráfica siguiente donde representamos los pares de datos (**peso, altura**)



- ¿Existe algún dato aberrante? ¿Podemos emplear esta gráfica para determinar una relación entre el peso y la altura de los individuos de esta población?

Vamos a construir ahora la variable

$$\mathbb{E}(\text{altura}|\text{peso} = y)$$

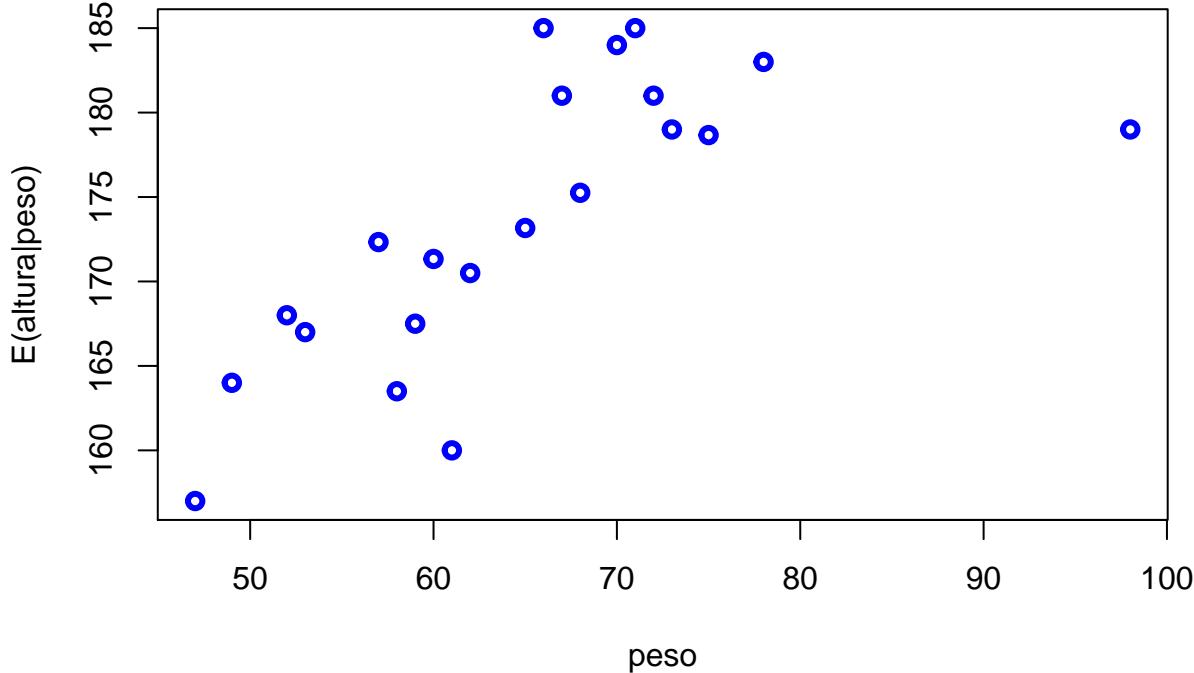
para todos los pesos y observados:

$$y \in \text{peso}(\Omega) = \{47, 49, 52, 53, 57, 58, 59, 60, 61, 62, 65, 66, 67, 68, 70, 71, 72, 73, 75, 78, 98\}$$

Obtenemos entonces una serie de datos

$$(\text{peso}, \mathbb{E}(\text{altura}|\text{peso}))$$

que podemos representar como



- Explica las diferencias entre este gráfico y el gráfico anterior. ¿Este gráfico mejora la representación de la relación entre el peso y la altura?

Procedemos a calcular los valores $\hat{\beta}$ y $\hat{\alpha}$ del modelo

$$\mathbb{E}(\text{altura}|\text{peso}) = \hat{\beta} \text{peso} + \hat{\alpha}$$

para ello empleamos la función de R `lm()`:

```
modelo_lineal <- lm(altura~peso)
summary(modelo_lineal)
```

```
##
## Call:
## lm(formula = altura ~ peso)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -14.550  -5.270   1.149   5.170  15.170 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 137.2105    8.0702   17.00 < 2e-16 ***
## peso         0.5700    0.1229    4.64 3.26e-05 ***
## ---
```

```
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.297 on 43 degrees of freedom
## Multiple R-squared:  0.3336, Adjusted R-squared:  0.3181
## F-statistic: 21.53 on 1 and 43 DF,  p-value: 3.264e-05
```

Obteniendo unos valores de $\hat{\beta} = 0.5699995$ y $\hat{\alpha} = 137.2104772$.

- ¿En qué unidades físicas se miden los parámetros del modelo $\hat{\beta}$ y $\hat{\alpha}$?
- Observamos que un individuo de esta población tiene un peso de 76 kg empleando el modelo construido enuncia que tipo de predicción se puede hacer acerca de la altura de dicho individuo.