

Aprendizaje Automatico - Trabajo Practico 3

Gonzalo Castiglione - 49138

May 10, 2012

Objetivo: Aplicar diversos métodos estadísticos para aprender a hacer inferencia a partir de datos experimentales.

1 Métodos de estadística paramétrica

1. Soluciones

(a) medidas:

	Ancho Pétalo	Largo Pétalo	Ancho Cépalo	Largo Cépalo
(b) $\hat{\mu}$	6.5880	2.9740	5.5520	2.0260
$\hat{\sigma}$	0.6232	0.3160	0.5409	0.2692
E_{cm}	0.0081	0.0021	0.0061	0.0015

(c) Intervalos de confianza para con un nivel de confianza de 0.95.

	Ancho	Largo	Ancho	Largo
I	2.8643	1.9657	3.0837	2.0863

2. Se tienen 80 componentes, de las cuales 12 son defectuosas.

(a) La proporción de componentes no defectuosos de la muestra $= \bar{x}_{nd} \frac{80-12}{80} = 0.85$

i. Un estimador \hat{x} es un estimador insesgado para estimar a x si $E[\hat{x}] = x$. Por lo tanto este es un estimador *insesgado*.

ii. Muestra: 68 mediciones con $\{x_i, y_i\} = 1$ y 12 mediciones con $\{x_i, y_i\} = 0$.

$e_0 = (0 - 0.85)$ (para las 12 muestras defectuosas)

$e_1 = (1 - 0.85)$ (para las 68 muestras no defectuosas)

Por lo que el error cuadrático medio estaría dado por la fórmula:

$$E_{CM} = \sqrt{\frac{(1-0.85)^2 * 68 + (0-0.85)^2 * 12}{80}} = \sqrt{\frac{1.53 + 8.67}{80}} \simeq 0.35$$

(b) Proporción de sistemas que funcionan correctamente $= \frac{\binom{80-12}{2}}{\binom{80}{2}} = \frac{2278}{3160} = 0.72$

3. código:

```
birds=[
...replace meditations here...
];
alpha=0.01
[h,p,ci,stats] = ttest(birds(:,2), birds(:,3), alpha)
```

Si $\alpha > p\text{-value} = 0.0117$

Entonces se rechaza H_0 (o sea que $\mu_D=0$)

Si no

Entonces no se rechaza, es decir las plumas podrían ser iguales.

Están en lo cierto, para $\alpha = 0.05$ hay variación entre el color de las plumas

4. Solución

(a) Gráfico del peso del cerebro y peso total para cada ejemplo dado

i. código

```
load brains.txt;
x = 1:28;
y = brains(:,1);
z = brains(:,2);
clf;
hold on;
plot(x, y, '*b;Peso Promedio en Kg;')
plot(x, z, '*r;Peso Cerebro Promedio en G;')
print('-dpng', './TotalWeightVsBrainWeight.png');
```

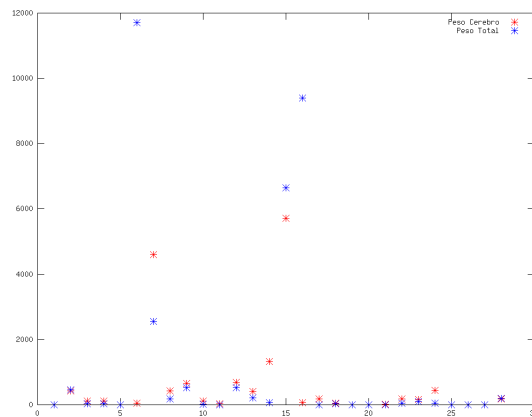


Figure 1: Peso del cerebro y peso total para cada medición en brains.txt*

* El valor del peso del cerebro de la medición 25 no se ve en la figura ya que se aleja demasiado del resto de los valores y el ajustar los ejes solo para mostrar ese valor produce que todas las demás mediciones no puedan apreciarse correctamente.

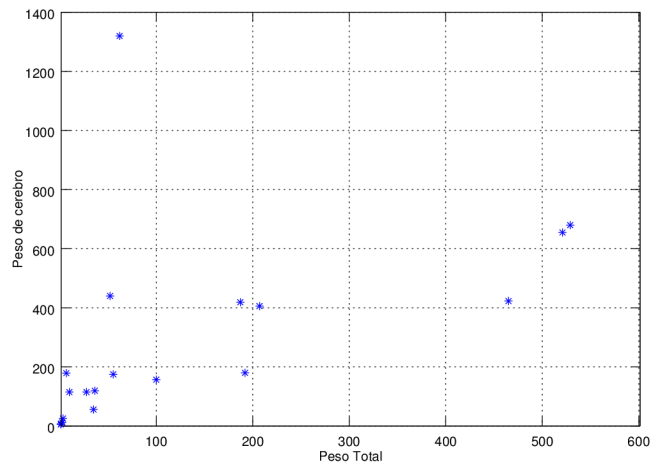


Figure 2: Peso total en Kg Vs peso del cerebro en G**

** Se removieron los valores para los 4 valores de x mayores a 2000 ya que ocultaban la visualización de todos los demás valores

En una observación a simple vista, se puede ver que las mediciones que se diferencian notablemente del resto son: 6, 7, 14, 15, 16 y por supuesto, la 25.

ii. No

(b) código

```
load brains.txt
regstats(brains(:,1), brains(:,2), 'linear')
```

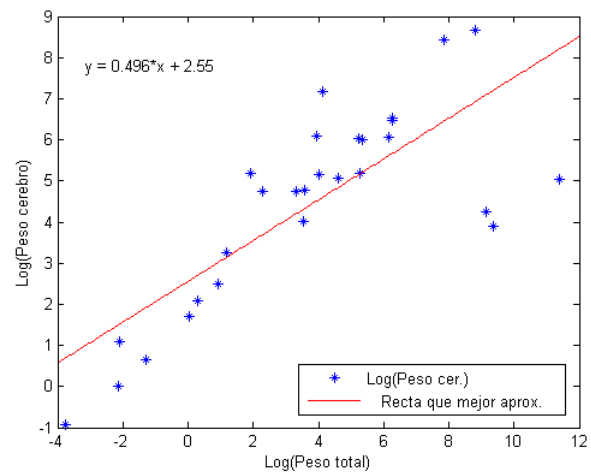


Figure 3: Logaritmo de ambas mediciones y la línea que mejor los aproxima