Ejercicio 1.2 de los cangrejos:

Realizaremos el estudio de una población de 200 cangrejos en de acuerdo a su especie, sexo y distintas longitudes de ciertas partes de los cangrejos

Al realizar un análisis descriptivo a través de un grafico (plot) y comparando sus cinco variables

(’FL’ frontal lip size’RW’ rear width ’CL’ carapace length ’CW’ carapace width ’BD’ body depth ), podemos ver que existe una relacion muy directa de acuerdo al sexo o a la especie.

(grafico de todas las variables comparadas)

Pero con estos datos y estas relaciones no es posible determinar el sexo o especie de un cangrejo si solo tenemos una de las 5 variables cuantitativas.

A continuación ejemplos de comparaciones en base al sexo o la especie para distintas variables cuantitativas:

(Grafico de 2 variables comparadas con el sexo o la especie)

Si observamos bien las comparaciones anteriores, pareciera como si fuese posible realizar una discriminación de acuerdo a la especie o al sexo del cangrejo de acuerdo a alguna de las variables cuantitativas.

Pero, si hacemos un análisis matemático mas profundo, y realizamos un estudio de correlacion de las variables cuantitavias con la función **cor()** podemos observar que todas las variables cuantitativas, tienen una gran correlacion entre si (superior al 0.85), por lo cual esto indicaría que si tenemos cangrejos con lobulos frontales grandes, estos tienen una gran probabilidad de tener pinzas pequeñas, o si tenemos un cangrejo con caparazón grande, también puede llegar a tener un caparazón angosto y pinzas de adulto, por lo que sabremos que no sera muy conveniente juntar comparar todas las variables cuantitativas al mismo tiempo ya que los resultados no serian muy coherentes.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | FL | RW | CL | CW | BD |
| FL | 1.0000000 | 0.9069876 | 0.9788418 | 0.9649558 | 0.9876272 |
| RW | 0.9069876 | 1.0000000 | 0.8927430 | 0.9004021 | 0.8892054 |
| CL | 0.9788418 | 0.8927430 | 1.0000000 | 0.9950225 | 0.9832038 |
| CW | 0.9649558 | 0.9004021 | 0.9950225 | 1.0000000 | 0.9678117 |
| BD | 0.9876272 | 0.8892054 | 0.9832038 | 0.9678117 | 1.0000000 |

Una forma de obtener resultados mas coherentes y mas realistas es tratando los datos iniciales y centralizandolos para poder analizarlos de una manera mas correcta. Para lograr esto, primero debemos sumar cada valor de cada variable cuantitativa por especie (es decir, para el cangrejo 1, sumar el valor de cada variable cuantitativa) y luego esta dividirla por cada valor individual de las variables cuantitativas (ejemplo al valor de la variable cuantitiva FL del cangrejo 1, hay que dividirla por el valor de la sumatoria realizada anteriormente), y si luego de esto, aplicamos la función **cor()**, podemos observar que los resultados de la matriz cambian notoriamente:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | FL | RW | CL | CW | BD |
| FL | 1.0000000 | -0.1183474 | -0.24026563 | -0.7825549 | 0.51974244 |
| RW | -0.1183474 | 1.0000000 | -0.82661717 | -0.2020264 | -0.46394103 |
| CL | -0.2402656 | -0.8266172 | 1.00000000 | 0.4186543 | 0.09676383 |
| CW | -0.7825549 | -0.2020264 | 0.41865431 | 1.0000000 | -0.64982922 |
| BD | 0.5197424 | -0.4639410 | 0.09676383 | -0.6498292 | 1.00000000 |

Ejercicio 2.3 de los cangrejos

En primera instancia y al utilizar la función princomp() sobre la variable crabsquant, podemos observar que se manifiesta un componente principal (Comp 1) con una desviación estándar superior a los demás componentes principales.

*(Dans une première instance et après avoir utilisé la fonction princomp () sur le variable crabsquant, nous pouvons observer que se manifeste un composant principal (Comp 1) avec une déviation standard supérieure aux autres composants principaux.)*

Standard deviations:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

11.8322521 1.1359369 0.9976311 0.3669098 0.2784325

Y si esto lo graficamos con la función biplot(), podemos ver claramente que todas las variables (’FL’ frontal lip size’RW’ rear width ’CL’ carapace length ’CW’ carapace width ’BD’ body depth ) estan orientadas hacia el primer componente. Lo cual puede quedar mejor reflejado en esta tabla de componentes principales:

*(Et si cela le graficamos avec la fonction biplot (), nous pouvons voir clairement que toutes les variables ('FL' frontal lip size'RW ' rear width ' CL ' carapace length ' CW ' carapace width ' BD ' body depth) sont orientées vers le premier composant. Ce qui peut rester meilleur reflété dans cette table de composants principaux :)*

PC1 PC2 PC3 PC4 PC5

FL 0.2889810 0.3232500 -0.5071698 0.7342907 0.1248816

RW 0.1972824 0.8647159 0.4141356 -0.1483092 -0.1408623

CL 0.5993986 -0.1982263 -0.1753299 -0.1435941 -0.7416656

CW 0.6616550 -0.2879790 0.4913755 0.1256282 0.4712202

BD 0.2837317 0.1598447 -0.5468821 -0.6343657 0.4386868

El problema de este análisis, es que no nos permite hacer un estudio ni una distinción de los cangrejos en cuanto al sexo o color que tiene cada cangrejo.

*(Le problème de cette analyse, c'est qu'il(elle) ne nous permet pas de faire d'étude ni une distinction des crabes en ce qui concerne le sexe ou la couleur qui a chaque crabe.)*

*Respuesta b*

*Ya que el analisis anterior no puede realizar un buen estudio respecto a los cangrejos analizados, es necesario realizar un tratamiento sobre los datos y realizar una distinción entre sexo y especie.*

*Para ello primero debemos centralizar la tabla, dividiendo cada variable, por la suma de de las variables y luego obtener los ACP correspondientes.*

*(mostrar graficos)*

*Como se puede ver en los graficos, existe una mejora notable en cuanto a la distinción de los cangrejos analizados, diferenciando de manera correcta a que grupo pertenece cada cangrejo y dando una impresión que es posible identificar a que grupo pertence cada cangrejo estudiado de acuerdo al grupo y a la especie.*