

Documentație proiect

Cerinta 1

Linia 18, 34, 50: *mdn = median(edu, na.rm = FALSE)*

```
18 mdn = median(edu, na.rm = FALSE)
```

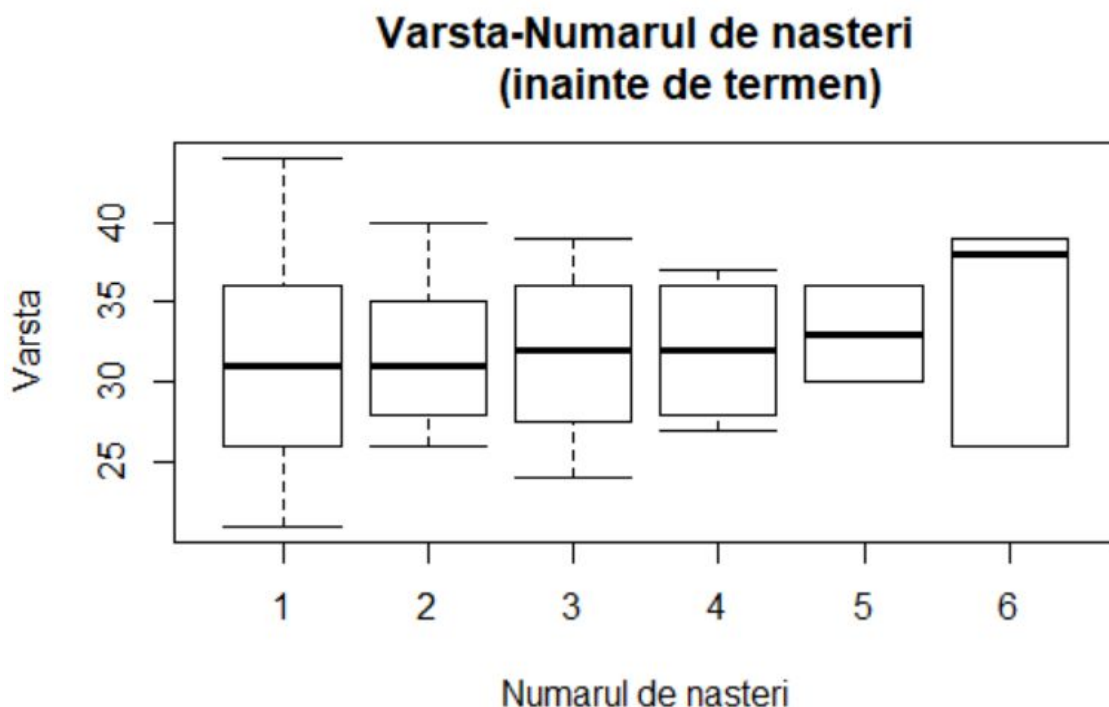
```
34 mdn = median(varsta, na.rm = FALSE)
```

```
50 mdn = median(paritate, na.rm = FALSE)
```

Se poate observa ca mediana este egală cu a doua quartilă.

Liniile 53 - 69:

```
53 varstasrt = sort(varsta)
54 varsta
55 varstasrt
56 lg<-length(varstasrt)
57 boxplot(varsta~paritate,data=infert, main="Varsta-Numarul de nasteri
58         (inainte de termen)",
59         xlab="Numarul de nasteri", ylab="Varsta", staplewex = 1, ylim=c(varstasrt[1], varstasrt[lg]))
60
61
62 boxplot(varsta~paritate,data=infert, main="Varsta-Numarul de nasteri(inainte de termen)",
63         xlab="Numarul de nasteri", ylab="Varsta", staplewex = 1, ylim=c(varstasrt[1], varstasrt[lg]))
64
65 text(y=fivenum(varsta~paritate), labels =fivenum(varsta~paritate), x=0.25)
66
67 boxplot(varsta, horizontal = TRUE, axes = FALSE, staplewex = 1)
68 text(x=fivenum(X), labels =fivenum(X), y=1.25)
69
```



Pentru seturile de date folosite în cadrul analizei de boxplot, observăm modul în care este repartizată variabila Paritate în funcție de variabila Vârstă.

Linia orizontală din mijloc reprezintă mediana (quartila 2).

Quartilele 1 și 3 delimitează capetele boxplot-ului.



Mustașa superioară este determinată de valoarea de celei mai mari observații, care este $\leq q3 + 1.5 \cdot (q3 - q1)$.

Mustașa inferioară reprezintă valoarea celei mai mici observații (adică cea mai mică vârstă), și este $\geq q1 - 1.5 \cdot (q3 - q1)$.

Valorile din afara dreptunghiului se numesc valori aberante (extreme).

Boxplot-ul 6 nu are valori extreme, deoarece quantilele sale sunt:

```
vec <- c(26,26,26,38,38,39,39,39)
```

Mediana pentru boxplot-ul 6 arată o compactare a 1/2 din date în jurul intervalului [38-39], în raport cu mai marele rămas sub mediană [26-38].

Quartila 1 dă cele mai mici elemente, iar quartila 3 este egală cu quartila 4 și nu avem valori extreme (ieșite din box).

Linia 90, 106, 122: `mdn = median(varsta, na.rm = FALSE)`

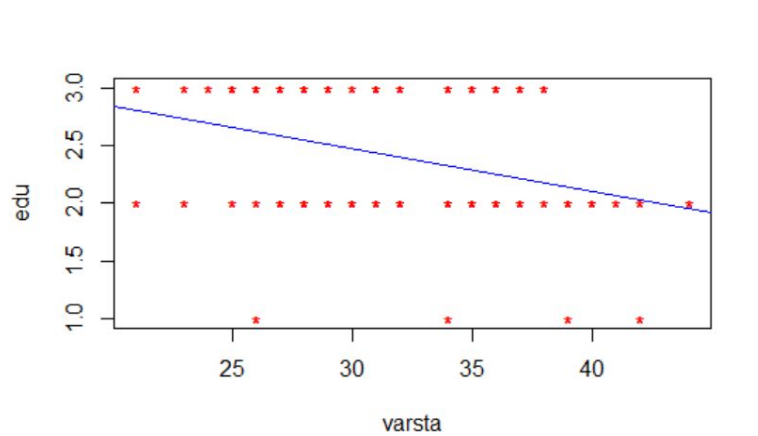
```
90 mdn = median(indus, na.rm = FALSE)
```

```
106 mdn = median(case, na.rm = FALSE)
```

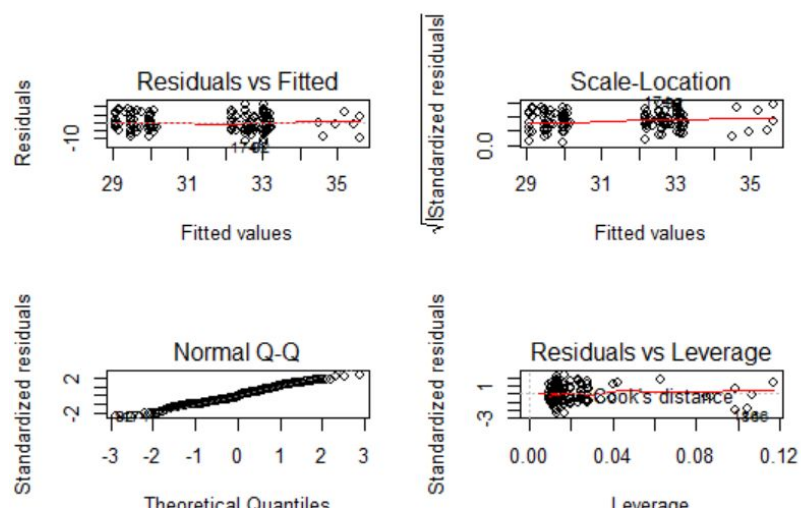
```
122 mdn = median(spontan, na.rm = FALSE)
```

Se poate observa că mediana este egală cu a doua quartilă.

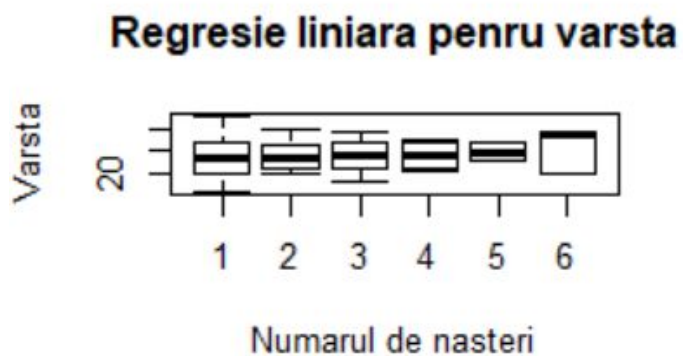
Cerința 2: Regresie simplă



Cerința 2: Regresie multiplă



Cerința 2: Regresie liniară simplă cu o variabilă independentă



Liniile 196 - 204:

```
196 # Regresie liniara multipla
197 # Regresie liniara multipla cu 2 variabila independente prin metoda celor mai mici patrate
198 varsta <- model$age #set1 de date initial -> selectie de regresie liniara
199 length(varsta)
200 paritate<-(model$parity) #set2 de date initial -> selectie de regresie liniara
201 c=cbind(1,varsta,paritate)
202 yrage=(1/4+paritate+varsta) #set combinat al datelor initiale
203 round(yrage, digits = 0)
204 solve(t(c) %*% c, t(c) %*% yrage)
```

În boxplot se poate observa că vârstele au fost micșorate cu 25% și că s-a calculat și numărul de nașteri corespunzător.

Interpretare cerința 2:

beta0= 0.25

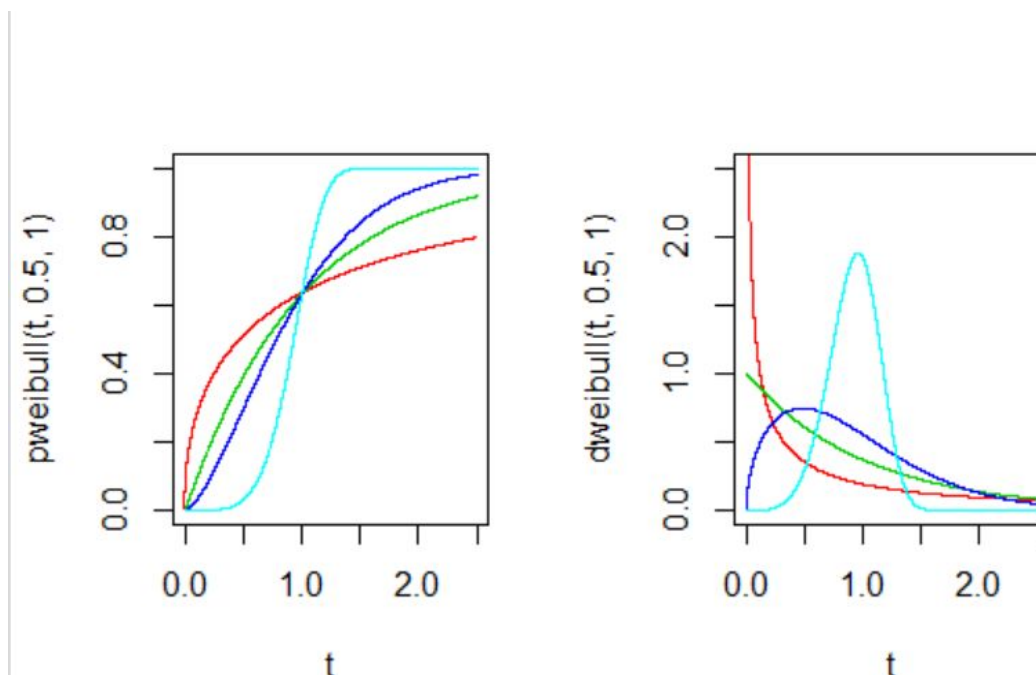
beta1= 1

alfa= -3.352547

beta=7.500000

Construcția setului de date infer, fiind format din mai multe variabile aleatoare, îi dau posibilitatea de a fi folosit cu regresie liniară multiplă. Varianta regresie multiplă e mai potrivită decat cea simplă.

Distribuția Weibull



În prima diagramă sunt afișate diferite funcții repartizate Weibull, dar cu valori diferite ale parametrilor.

În diagrama a doua este prezentată funcția densitate de repartiție pentru fiecare funcție din diagrama anterioară.

Valoarea maximă ce poate fi atinsă de funcțiile din prima diagrama este 1.

Repartitia Weibull este folosită pentru:

- prezicerea schimbărilor meteorologice, a evenimentelor extreme precum maximă anuală de precipitații a unei zile, a prezice debitul unui râu în hidrologie, a descrie distribuția vânturilor
- strângerea de informații de pe un site web - modelează timpul de căutare pe un site
- asigurări generale, pentru a prezice revendicările de reasigurare