

MAT 2779, Introduction à la biostatistique

Devoir 2
SOLUTIONNAIRE

Professeur : M'hammed Mountassir

UN TOTAL DE 20 POINTS

Question 1 : Problème 4.2 du livre (5 points) .

Soit X le nombre de bébés avec une anomalie congénitale dans l'échantillon . Alors X suit une loi binomiale avec $n = 100$ et $p = 1/25$.

(a) On veut calculer

$$P(X \leq 2) = (24/25)^{100} + 100(1/25)^1(24/25)^{99} + 4950(1/25)^2(24/25)^{98} = 0.2321.$$

Qu'on peut aussi calculer avec le logiciel R, en tapant `pbinom(2,100,1/25)` ce qui donne : 0.2321

(b) La moyenne de X est : $E(X) = np = (100)(1/25) = 4$. (c) L'écart type de X est : $\sigma = \sqrt{np(1-p)} = \sqrt{100(1/25)(24/25)} = 1.9596$.

Question 2 : Problème 5.2 du livre (5 points).

Soit X la température les 15 Janvier ,

$$P(-20 < X < -8) = p$$

se calcule avec le logiciel R, en tapant : `pnorm(-8,-12,5)-pnorm(-20,-12,5)` ce qui donne $p = 0,7333$ Soit Y le nombre d'années les grappes sont cueillis un 15 Janvier pendant 5 années consécutives, alors cette variable suit une loi binomiale de paramètres $n = 5$ et $p = 0.7333$ on a alors :

$$P(Y = 5) = \frac{5!}{5! 0!} (0.7333)^5 (1 - 0.7333)^0 = (0.7333)^5 = 0.2120$$

. Qu'on peut aussi calculer avec R.

Question 3 : Problème 5.6 du livre (5 points).

- (a) Vraie ;
- (b) Fausse ;
- (c) Vraie : On a

$$0.75 = P(X < x) = P\left(Z < \frac{x - 5}{2}\right)$$

par conséquent $(x - 5)/2 = \text{qnorm}(0.75, 0, 1)$, i.e. $x = 2 * \text{qnorm}(0.75, 0, 1) + 5$;

- (d) fausse :

Question 4 : Problème 7.4 du livre (5 points)

- (a) La moyenne est égale à $\bar{x} = 1.99$ et l'écart type est $s = 0.661$.
- (b) Si on met ces données dans un vecteur, le logiciel R nous donne : $Q2 = 2,50$ $Q1 = 1,525$; $Q3 = 2,375$ et $DIQ = Q3 - Q1 = 0,85$
- (c) Pour trouver les données aberrantes éventuelles on calcule :

$$\text{Lim}_{\text{inf}1} = q_1 - (1.5)DIQ = 1.525 - (1.5)(0,85) = -0.25$$

$$\text{Lim}_{\text{sup}2} = q_3 + (1.5)DIQ = 2.375 + (1.5)(0.85) = 3,65$$

Il n'y a aucune données à l'extérieur de ces limites, donc il n'y a pas de donnée aberrante.

On peut aussi seulement faire un boxplot avec le logiciel R.

Question 5 : (À NE PAS CORRIGER) Utiliser le logiciel R

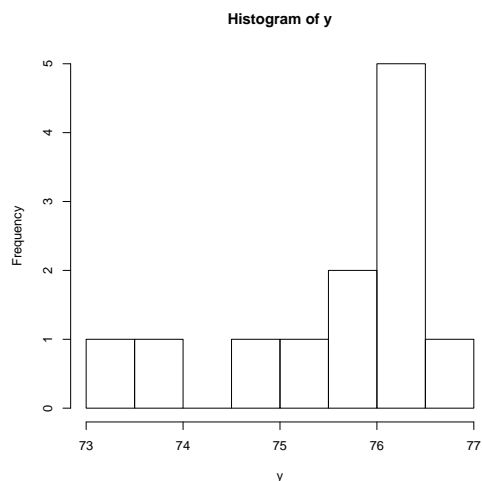
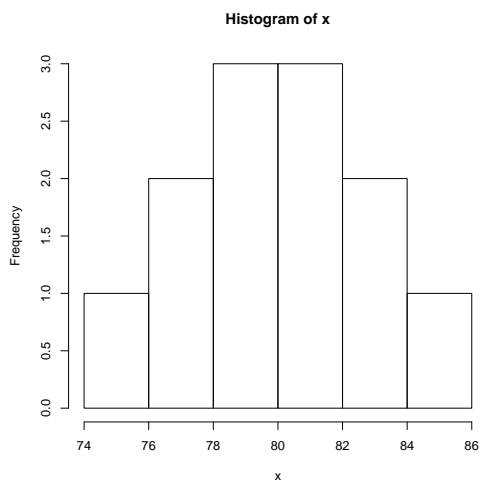
On met les données dans deux vecteurs x pour l'espèce 1 et y pour l'espèce 2.

```
> summary(x)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
75.70  78.07   79.70   79.89   81.88   84.20
> summary(y)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
73.10  75.35   76.05   75.61   76.30   76.80
```

b) Les histogrammes de x et y sont les suivants :

```
> hist(x, breaks=6)
```

```
> hist(y, breaks=6)
```



La distribution de x est symétrique. La distribution de y a une assymétrie vers la gauche .

c) On peut aussi donner un `boxplot(x,y)`.