



Corrigé de l'examen CT de statistiques

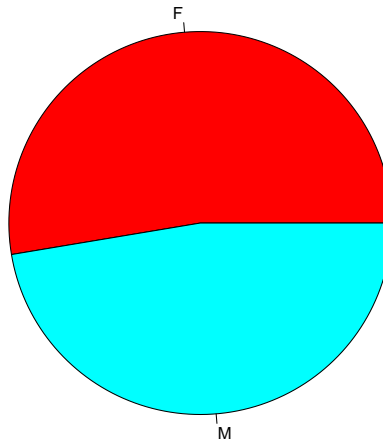
On étudie le fichier de données 'M1IGAPASA10data.txt'.

- (1) (a) • On étudie la variable qualitative (ou catégorielle) 'sexe'.
 - Les effectifs et les pourcentages déterminés par \mathcal{R} sont donnés dans le tableau suivant

	effectifs	pourcentages
F	10	52.632
M	9	47.368

•

Camembert pour sexe

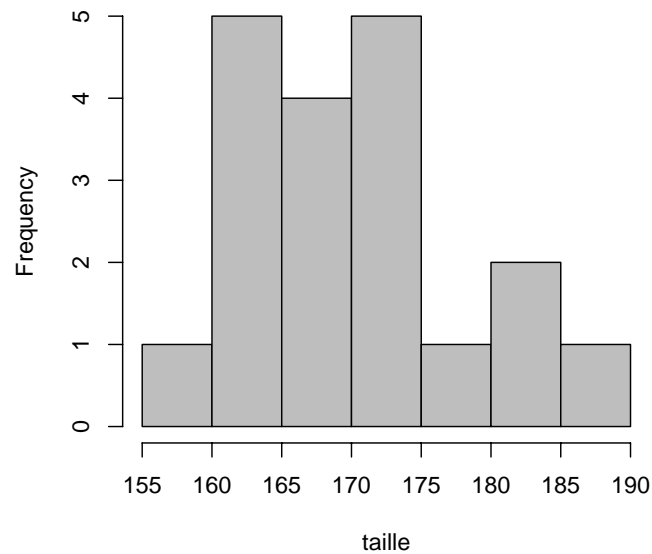


- (b) • On étudie la variable quantitative (ou numérique) 'taille'. Pour les manipulations avec \mathcal{R} , on renvoie donc à la section 3.4 et aux sections récapitulatives 7.1.1 et 7.1.3 du document de cours.
- Les différents résultats déterminés par \mathcal{R} sont donnés dans le tableau suivant

noms	valeurs
moyenne	171.05
écart-type	8.76
Q_1 (quartile à 25 %)	164
médiane	170
Q_3 (quartile à 75 %)	174.5
minimum	157
maximum	190
nombre	19

•

Histogramme pour taille

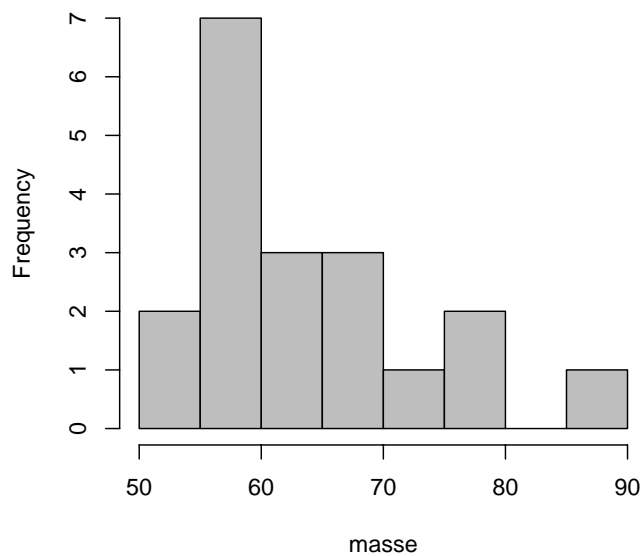


- (c) • On étudie la variable quantitative (ou numérique) 'masse'. Pour les manipulations avec \mathcal{R} , on renvoie donc à la section 3.4 et aux sections récapitulatives 7.1.1 et 7.1.3 du document de cours.
- Les différents résultats déterminés par \mathcal{R} sont donnés dans le tableau suivant

noms	valeurs
moyenne	64.53
écart-type	10.1
Q_1 (quartile à 25 %)	57
médiane	63
Q_3 (quartile à 75 %)	67.5
minimum	52
maximum	90
nombre	19

•

Histogramme pour masse

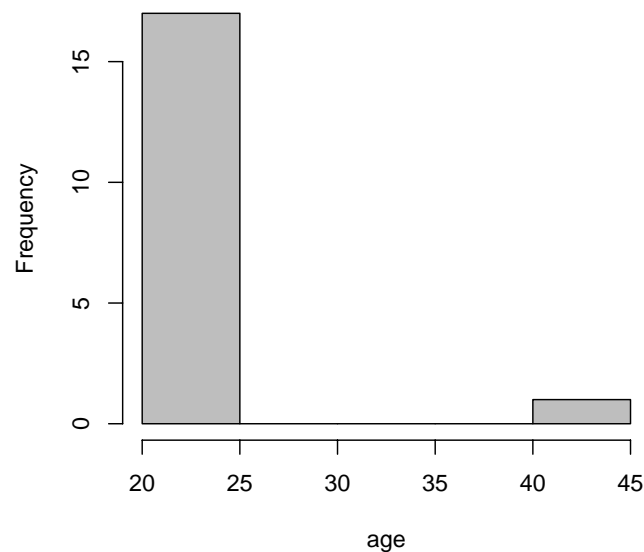


- (d) • On étudie la variable quantitative (ou numérique) 'age'. Pour les manipulations avec \mathbb{R} , on renvoie donc à la section 3.4 et aux sections récapitulatives 7.1.1 et 7.1.3 du document de cours.
- Les différents résultats déterminés par \mathbb{R} sont donnés dans le tableau suivant

noms	valeurs
moyenne	22.94
écart-type	5.41
Q_1 (quartile à 25 %)	21
médiane	21
Q_3 (quartile à 75 %)	23
minimum	20
maximum	44
NA	1
nombre	19

•

Histogramme pour age



- (2) (a) • On étudie le croisement de la variable qualitative (ou catégorielle) 'sexe' et de la variable quantitative (ou numérique) 'age'. Pour les manipulations avec \mathbb{R} , on renvoie donc aux sections 6.2 et 6.3 et la section récapitulative 7.2.3 du document de cours.
- Avec \mathbb{R} , on obtient les statistiques par groupes données dans le tableau suivant ;

	mean	sd	0%	25%	50%	75%	100%	n	NAs
F	21.778	1.481	20.000	21.000	21.000	22.000	25.000	9	1
M	24.111	7.541	20.000	21.000	21.000	23.000	44.000	9	0

On rappelle que, dans ce tableau :

- le nombre noté 0% est le quartile à 0 % (c'est le minimum) ;
- le nombre noté 25% est le quartile à 25 % (c'est Q_1) ;
- le nombre noté 50% est le quartile à 50 % (c'est la médiane) ;
- le nombre noté 75% est le quartile à 75 % (c'est Q_3) ;
- le nombre noté 100% est le quartile à 100 % (c'est le maximum).

Les autres résultats donnés par \mathbb{R} sont les suivants :

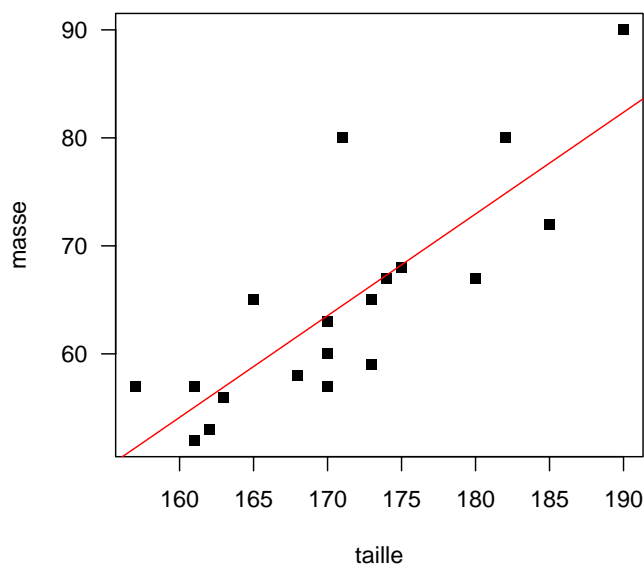
Noms des indicateurs	Valeurs
Rapport de corrélation RC	0.049301
probabilité critique p_c	0.375871

On compare le rapport de corrélation $RC=0.049301$ aux seuils de Cohen (0.01,0.05,0.15) (voir [Coh92]) et la probabilité critique $p_c=0.375871$ à la valeur seuil de la probabilité critique 0.05 et on déduit les résultats suivants sur la significativité de la liaison :

significativité pratique	moyenne
significativité statistique	non

- (b) • On étudie le croisement de la variable quantitative (ou numérique) 'taille' et de la variable quantitative (ou numérique) 'masse'. Pour les manipulations avec \mathbb{R} , on renvoie donc à la section 4.5 et la section récapitulative 7.2.1 du document de cours.
- Voir la figure ci-dessous.

masse en fonction de taille



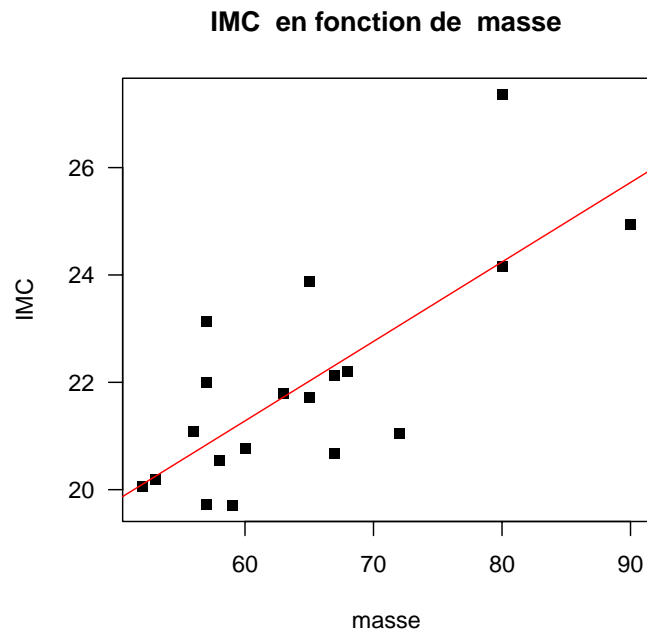
- Les résultats donnés par \mathbb{R} sont les suivants :

Noms des indicateurs	Valeurs
pente a	0.941726
ordonnée à l'origine b	-96.55835
corrélation linéaire r	0.816563
probabilité critique p_c	1.99893e-05

On compare la valeur absolue de la corrélation linéaire $r = 0.816563$ aux seuils de Cohen (0.1,0.3,0.5) (voir [Coh92]) et la probabilité critique $p_c = 1.99893e-05$ à la valeur seuil de la probabilité critique 0.05 et on déduit les résultats suivants sur la significativité de la liaison linéaire :

significativité pratique	très forte
significativité statistique	oui

- (3) (a) • On étudie le croisement de la variable quantitative (ou numérique) 'masse' et de la variable quantitative (ou numérique) 'IMC'.
 • Voir la figure ci-dessous.



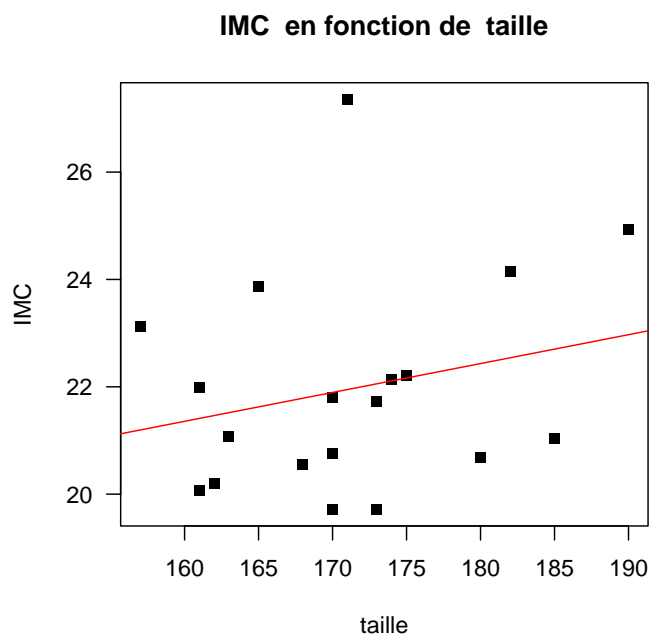
- Les résultats donnés par \mathcal{R} sont les suivants :

Noms des indicateurs	Valeurs
pente a	0.148043
ordonnée à l'origine b	12.39885
corrélation linéaire r	0.750646
probabilité critique p_c	0.000213154

On compare la valeur absolue de la corrélation linéaire $r = 0.750646$ aux seuils de Cohen (0.1,0.3,0.5) (voir [Coh92]) et la probabilité critique $p_c = 0.000213154$ à la valeur seuil de la probabilité critique 0.05 et on déduit les résultats suivants sur la significativité de la liaison linéaire :

significativité pratique	très forte
significativité statistique	oui

- (b) • On étudie le croisement de la variable quantitative (ou numérique) 'taille' et de la variable quantitative (ou numérique) 'IMC'.
- Voir la figure ci-dessous.



- Les résultats donnés par \mathcal{R} sont les suivants :

Noms des indicateurs	Valeurs
penne a	0.053797
ordonnée à l'origine b	12.749436
corrélation linéaire r	0.236521
probabilité critique p_c	0.329607

On compare la valeur absolue de la corrélation linéaire $r = 0.236521$ aux seuils de Cohen (0.1,0.3,0.5) (voir [Coh92]) et la probabilité critique $p_c = 0.329607$ à la valeur seuil de la probabilité critique 0.05 et on déduit les résultats suivants sur la significativité de la liaison linéaire :

significativité pratique	moyenne
significativité statistique	non

Références

[Coh92] J Cohen. A power primer. *Psychological bulletin*, 112(1) :155–159, 1992.