NOM : Date : .

PRENOM : Groupe : .

Mathématiques pour la Biologie : Feuille-réponses du TD 8
Représentation de variables statistiques à 1 et 2 dimensions

Exercice 1. (moyenne, variance) : Pour déterminer l'effet de nutriment sur la croissance de plants de haricots, on prépare deux jeux de 100 pots de terre, les premiers contenant un apport de nutriment (pots expérimentaux) et les seconds n'en contenant pas (pots de contrôle). On plante un haricot dans chaque pot et on mesure, après 20 jours, la taille des plants obtenus. Voici les 6 premières mesures de chaque lot (pour simplifier l'exercice, on ne prendra en compte que ces 6 mesures).

Contrôle	9,3	13,2	14,2	11,7	10,3	9,8
Expérimental	13,4	12,1	15,2	13,2	13,5	14,9

1. Calculer la moyenne μ_c et μ_e de chaque lot.

$\mu_c = (9,3 + 13,2 + 14,2 + 11,7 + 10,3 + 9,8)/6 = 68,5/6 \simeq 11,42$

$\mu_e = (13,4 + 12,1 + 15,2 + 13,2 + 13,5 + 14,9)/6 = 82,3/6 \simeq 13,72$

2. Calculer la variance Var_c et l'écart-type σ_c du lot de contrôle. Même question pour l'autre lot.

$\mathrm{Var}_c = \frac{1}{6}\left((9,3-\mu_c)^2 + (13,2-\mu_c)^2 + \cdots + (9,8-\mu_c)^2\right) \simeq 3,225$ d'où $\sigma_c = \sqrt{\mathrm{Var}_c} \simeq 1,80$

$\mathrm{Var}_e = \frac{1}{6}\left((13,4-\mu_e)^2 + (12,1-\mu_e)^2 + \cdots + (14,9-\mu_e)^2\right) \simeq 1,105$ d'où $\sigma_e = \sqrt{\mathrm{Var}_e} \simeq 1,05$

3. Pour chaque lot, indiquer combien de mesures sont dans l'intervalle $[\mu - \sigma, \mu + \sigma]$. Même question pour l'intervalle $[\mu - 2\sigma, \mu + 2\sigma]$.

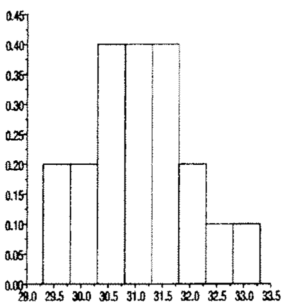
$[\mu_c - \sigma_c, \mu_c + \sigma_c] = [9,6 ; 13,2]$ qui contient $\boxed{4}$ données sur les 6

$[\mu_c - 2\sigma_c, \mu_c + 2\sigma_c] = [7,8 ; 15]$ qui contient toutes les 6 données

$[\mu_c - \sigma_e, \mu_e + \sigma_e] = [12,65 ; 14,75]$ qui contient $\boxed{3}$ données sur les 6

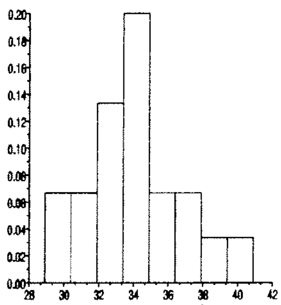
$[\mu_e - 2\sigma_e, \mu_e + 2\sigma_e] = [11,6 ; 15,8]$ qui contient toutes les 6 données

Exercice 2. (histogrammes) :



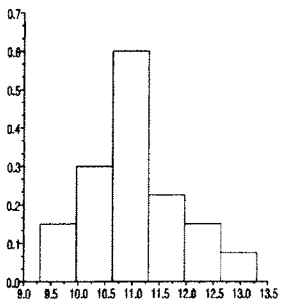
$\mu \simeq 31$ C

σ petit



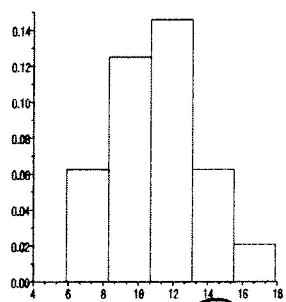
$\mu \simeq 34$ A

σ grand



$\mu \simeq 11$ B

σ petit



$\mu \simeq 11$ D

σ grand

Associer à chaque histogramme le bon jeu de paramètres parmi les 4 jeux suivants : A : ($\mu = 34,285, \sigma = 2,89$), B : ($\mu = 11,095, \sigma = 0,96$), C : ($\mu = 31,095, \sigma = 0,96$) et D : ($\mu = 11,285, \sigma = 2,89$) en indiquant sous chaque dessin les valeurs sélectionnées.

Exercice 3. (covariance de deux séries) : Compléter les deux tableaux suivants (dans la dernière colonne, mettre la moyenne) et en déduire les valeurs de $\mathrm{Cov}(x, y)$ et $\mathrm{Cov}(x, z)$.

						μ
x_i	1	2	3	5	7	3,6
y_i	4	3	2	1	1	2,2
x_iy_i	4	6	6	5	7	5,6

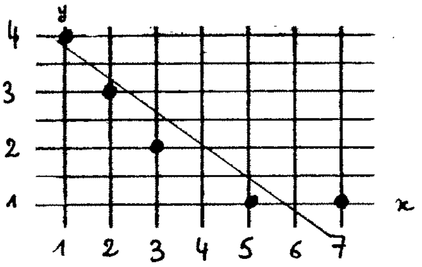
						μ
x_i	1	2	3	5	7	3,6
z_i	1	3	0	3	4	2,2
x_iz_i	1	6	0	15	28	10

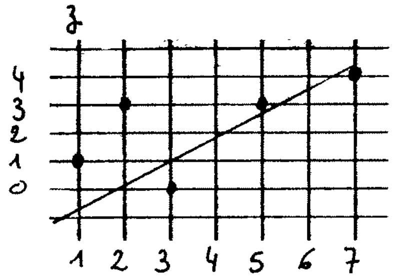
$\mathrm{Cov}(x,y) = 1/5 \sum x_iy_i - \mu_x\mu_y$
$= 5,6 - 3,6 \times 2,2 = -2,32$

$\mathrm{Cov}(x,z) = 1/5 \sum x_iz_i - \mu_x\mu_z$
$= 10 - 3,6 \times 2,2 = 2,08$

Tracer les deux nuages de points (x_i, y_i) et (x_i, z_i) et en déduire une justification des signes des covariances obtenues.

$\mathrm{Cov}(x,y) < 0$ le nuage est descendant





$\mathrm{Cov}(x,z) > 0$ le nuage est ascendant

Exercice 4. (transformation de la variance par une application affine) :

1. On considère la série des mesures de l'exercice 1 (que l'on note toujours x_i) et la série y_i qui en est une image par translation $y_i = x_i - \mu_c$ que l'on appelle série *centrée*. Compléter le tableau suivant et en déduire la variance de la série y_i.

							μ
x_i	9,3	13,2	14,2	11,7	10,3	9,8	11,42
$y_i = x_i - \mu_c$	-2,1	1,8	2,8	0,3	-1,1	-1,6	$\simeq 0$
y_i^2	4,41	3,24	7,84	0,09	1,21	2,56	3,225

la moyenne d'une série centrée est nulle mais les erreurs d'arrondi donnent 0,1

$\mathrm{Var}(y) = \frac{1}{6} \sum y_i^2 - \mu_y^2 \simeq 3,225 - 0^2 = 3,225$

2. Comparer avec la variance des x_i calculée précédemment. Quelle propriété de la variance explique ce résultat ? On avait trouvé $\mathrm{Var}(x) = \mathrm{Var}_c = 3,225$ donc $\mathrm{Var}(x) = \mathrm{Var}(y)$.
On a vu en cours que $\mathrm{Var}(ax+b) = a^2\,\mathrm{Var}(x)$, ici $a = 1$ et $b = -\mu_c$.
La variance ne change pas par translation.

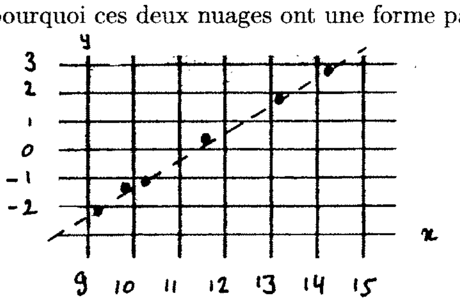
3. On transforme à présent la série x_i en la série $z_i = \frac{1}{2}x_i$. Calculer la variance des z_i et la comparer avec celle des x_i. Commentez.

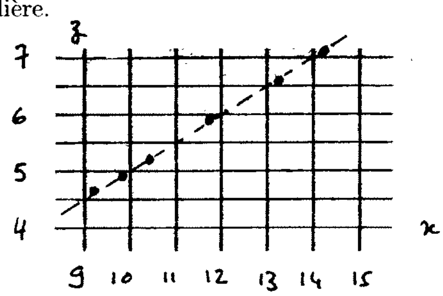
							μ
x_i	9,3	13,2	14,2	11,7	10,3	9,8	11,42
$z_i = \frac{1}{2}x_i$	4,65	6,6	7,1	5,85	5,15	4,9	5,71
z_i^2	21,62	43,56	50,41	34,22	26,52	24,01	33,39

$\mathrm{Var}(z) = \frac{1}{6} \sum z_i^2 - \mu_z^2 = 33,39 - (5,71)^2 \simeq 33,39 - 32,60 \simeq 0,79$

On applique $\mathrm{Var}(ax+b) = a^2\,\mathrm{Var}(x)$ avec $a = 1/2$ et $b = 0$: $\mathrm{Var}(z) = 0,79 \simeq (1/2)^2\, 3,225 \simeq 0,80$

4. Représenter les deux nuages de points (x_i, y_i) et (x_i, z_i) ainsi que leurs centres de gravité. Expliquer pourquoi ces deux nuages ont une forme particulière.





Les points (x_i, y_i) sont sur la droite $y = x - 11,42$

Les points (x_i, z_i) sont sur la droite $z = \frac{1}{2}x$