Correction de certains exercices de Statistiques

Cette correction vous permet de comparer vos résultats et de vous assurer de la justesse de vos raisonnements.
Dans la majorité des cas les calculs numériques ne sont pas détaillés.
Si vous repérez des erreurs, n'hésitez pas à m'en faire part.

Statistiques

Exercices du cours (11-statistiques descriptives cours & exos)

Exercice 1

(a) Pour qu'il y ait *progression*, il faut que le taux de réussite augmente. Hors le taux de réussite *global* passe de $\frac{15}{25}$ à $\frac{17}{25}$. Il y a donc une progression globale. L'affirmation est correcte.

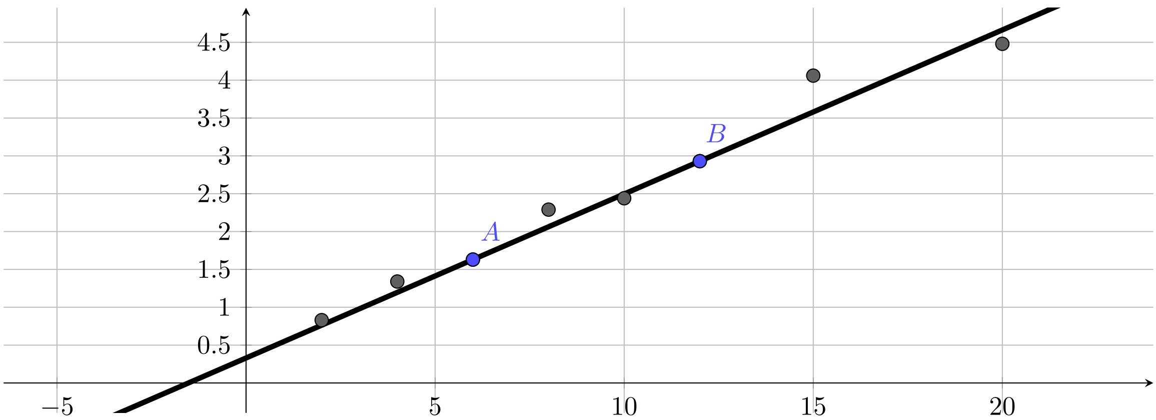
(b) Examinés en détail, les résultats font apparaître une régression à la fois chez les non-redoublants et chez les doublants. Néanmoins, cette affirmation est incorrecte, car imprécise. Elle contredit la précédente, qui est juste.

1. Le paradoxe tient à l'effet de structure. Si le taux de réussite des deux catégories diminue, l'effectif varie aussi considérablement.

   Sans cette variation d'effectif, il ne serait pas possible de diminuer les deux taux de réussite et d'obtenir une augmentation du taux de réussite.
2. La réponse précédente permet de résoudre le problème.

Exercice 2

1. Voici le graphe obtenu après la question 2.a)



2. (a) voir figure

   (b) On lit graphiquement pour $x = 18$ la valeur $y = 4.2$ On peut espérer une resistance de 4.2 m^2/w pour une épaisseur de 18 cm.
3. (a) La droite de Mayer est obtenue en séparant selon les x croissant le nuage en deux parties égales (à une valeur près.)

   Pour les 4 premières valeurs sur les x croissant on trouve le point moyen G_1 en calculant la moyenne des abscisses et la moyenne des ordonnées.

   On a $G_1(5, 1.53)$. On fait de même avec les quatre plus grandes valeurs pour x croissant et on trouve $G_2(14.25; 3.48)$.

   Tracer la droite de Mayer revient à construire ces deux points et les relier.

   Je m'abstiens pour ne pas surcharger le graphique.

   On détermine ensuite la droite de Mayer en donnant l'équation de cette droite $(G_1; G_2)$

   On utilise pour cela les formules $y = ax + b$ avec $a = \dfrac{y_b - y_a}{x_b - x_a}$ et $b = y_a - ax_a$.

   Il vient $a = \dfrac{3.48 - 1.53}{14.25 - 5} = 0.21$ et $b = 0.49$: $y = 0.21x + 0.49$

   (b) On applique le même principe avec $x = 18$ on obtient $y = 4.27$

   La différence est modeste.

   Remarque : Ces méthodes « graphiques » sont inutiles quand une machine fait les calculs à notre place. Ce sera toujours le cas dorénavant.

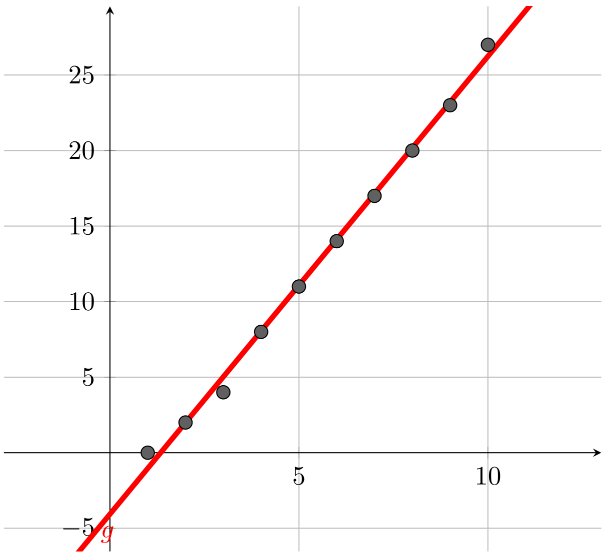
Exercice 3

1. A : 7, 8, 11, 12, 13, 13, 13 et B : 4, 7, 9, 12, 13, 13, 19
   mode : série A, la valeur la plus fréquente est 13. Série B le mode est 13.
   moyenne : série A : $\dfrac{7+8+11+12+13+13+13}{7} = 11$, série B : 11
   médiane : série A : 12 (3ème valeur), série B 12 (3ème valeur)
2. écart absolu moyen à la moyenne :
   On calcule les écarts à la moyenne, sans leurs signes et on fait la moyenne.
   Pour la série A ils valent : 4, 3, 0, 1, 2, 2, 2 et la moyenne est $\dfrac{14}{7} = 2$
   Inutile pour la série B, c'est stupide.
3. Variance on utilise la seule formule humainement acceptable : $V(X) = \overline{x^2} - \overline{x}^2$
   La méthode consiste à calculer les écarts précédents, les élever au carré et en prendre la moyenne.
   Pour la série A les carrés des écarts sont : 16, 9, 0, 1, 4, 4, 4 et la moyenne vaut $\dfrac{38}{7} = 5.43$.
   L'écart-type est la racine de la variance et vaut $\sigma(X) = \sqrt{V(X)} = 2.33$
   Pour la série B on trouve $V(X) = 11.22$ et $\sigma(X) = 3.35$

Par manque de temps, on passe directement à l'exercice 9.

Exercice 9

1. Voici le graphe obtenu après avoir tracé la droite de régression linéaire par la méthode des moindres carrés.



2. obtenir une droite de regression linéaire par la méthode des moindres carrés
   sur TI 83 : chercher (regression linéaire ti83), première vidéo :
   https ://www.youtube.com/watch ?v=uzMAL7uPE44
   sur casio : chercher (regression linéaire casio), première vidéo :
   https ://www.youtube.com/watch ?v=AYfsjiwsW00
   On obtient : $y = 3.03x - 4.07$
   Par manque de temps je zappe toute la recherche des coefficients...
   Seule la méthode à la calculatrice est au programme. Relisez le cours pour les détails.
3. Comme dans l'exercice précédent, on remplace x par 12 dans l'équation et
   $y = 3.03 \times 12 - 4.07 = 32.33$
   Au cours de 12 semestre d'utilisation, on estime qu'il y a 32.33% des lots qui vont subir une panne.