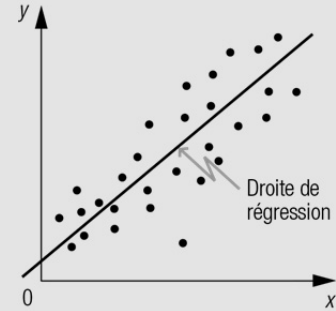


SAVOIRS

7.5 Droite de régression

- Dans un nuage de points mettant en relation deux variables statistiques, la droite qui représente le mieux l'ensemble des points du nuage est appelée **droite de régression**.
- La droite de régression permet de prédire la ou les valeurs de l'une des variables à partir des valeurs de l'autre.
- Le coefficient de corrélation linéaire permet de savoir jusqu'à quel point cette prédiction est fiable.

Exemple : Dans un nuage de points, il est possible de tracer une droite représentative de l'ensemble des points.

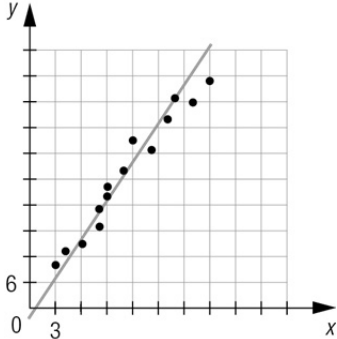


- Il existe diverses méthodes pour déterminer l'équation d'une droite de régression. En voici deux.

MÉTHODE DE LA DROITE DE MAYER

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|--|--|---|----------|----------|-----------------|----------|---------|----------|----------|----------|----------|----------|----|---|----|----|----|----|----|----|----|----|----|----|----|---|-----------------|--|--|--|--|--|-----------------|--|--|--|--|--|
| Démarche | <p>Exemple :</p> <table border="1" style="display: inline-table; margin-right: 10px;"> <tr><td>(21, 17)</td><td>(9, 40)</td><td>(15, 30)</td><td>(16, 25)</td><td>(26, 4)</td><td>(23, 14)</td></tr> <tr><td>(7, 45)</td><td>(25, 11)</td><td>(19, 20)</td><td>(11, 36)</td><td>(14, 34)</td><td>(18, 24)</td></tr> </table> | (21, 17) | (9, 40) | (15, 30) | (16, 25) | (26, 4) | (23, 14) | (7, 45) | (25, 11) | (19, 20) | (11, 36) | (14, 34) | (18, 24) | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (21, 17) | (9, 40) | (15, 30) | (16, 25) | (26, 4) | (23, 14) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (7, 45) | (25, 11) | (19, 20) | (11, 36) | (14, 34) | (18, 24) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1. Ordonner les couples de la distribution selon leurs abscisses. | <p style="text-align: center;">Distribution à deux variables</p> <table border="1" style="width: 100%; text-align: center;"> <tr><td>x</td><td>7</td><td>9</td><td>11</td><td>14</td><td>15</td><td>16</td><td>18</td><td>19</td><td>21</td><td>23</td><td>25</td><td>26</td></tr> <tr><td>y</td><td>45</td><td>40</td><td>36</td><td>34</td><td>30</td><td>25</td><td>24</td><td>20</td><td>17</td><td>14</td><td>11</td><td>4</td></tr> </table> | x | 7 | 9 | 11 | 14 | 15 | 16 | 18 | 19 | 21 | 23 | 25 | 26 | y | 45 | 40 | 36 | 34 | 30 | 25 | 24 | 20 | 17 | 14 | 11 | 4 | | | | | | | | | | | | |
| x | 7 | 9 | 11 | 14 | 15 | 16 | 18 | 19 | 21 | 23 | 25 | 26 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| y | 45 | 40 | 36 | 34 | 30 | 25 | 24 | 20 | 17 | 14 | 11 | 4 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2. Diviser la distribution en deux groupes, égaux si possible. | <table border="1" style="width: 100%; text-align: center;"> <tr><td>x</td><td>7</td><td>9</td><td>11</td><td>14</td><td>15</td><td>16</td><td>18</td><td>19</td><td>21</td><td>23</td><td>25</td><td>26</td></tr> <tr><td>y</td><td>45</td><td>40</td><td>36</td><td>34</td><td>30</td><td>25</td><td>24</td><td>20</td><td>17</td><td>14</td><td>11</td><td>4</td></tr> <tr><td colspan="6" style="text-align: center;">Groupe 1</td><td colspan="6" style="text-align: center;">Groupe 2</td></tr> </table> | x | 7 | 9 | 11 | 14 | 15 | 16 | 18 | 19 | 21 | 23 | 25 | 26 | y | 45 | 40 | 36 | 34 | 30 | 25 | 24 | 20 | 17 | 14 | 11 | 4 | Groupe 1 | | | | | | Groupe 2 | | | | | |
| x | 7 | 9 | 11 | 14 | 15 | 16 | 18 | 19 | 21 | 23 | 25 | 26 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| y | 45 | 40 | 36 | 34 | 30 | 25 | 24 | 20 | 17 | 14 | 11 | 4 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Groupe 1 | | | | | | Groupe 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3. Déterminer la moyenne des abscisses et la moyenne des ordonnées de chaque groupe pour former les couples moyens P ₁ (x ₁ , y ₁) et P ₂ (x ₂ , y ₂). | <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top; padding: 5px;"> $x_1 = \frac{7 + 9 + 11 + 14 + 15 + 16}{6} = 12$ $y_1 = \frac{45 + 40 + 36 + 34 + 30 + 25}{6} = 35$ <p style="text-align: center;">P₁(12, 35)</p> </td> <td style="width: 50%; vertical-align: top; padding: 5px;"> $x_2 = \frac{18 + 19 + 21 + 23 + 25 + 26}{6} = 22$ $y_2 = \frac{24 + 20 + 17 + 14 + 11 + 4}{6} = 15$ <p style="text-align: center;">P₂(22, 15)</p> </td> </tr> </table> | $x_1 = \frac{7 + 9 + 11 + 14 + 15 + 16}{6} = 12$ $y_1 = \frac{45 + 40 + 36 + 34 + 30 + 25}{6} = 35$ <p style="text-align: center;">P₁(12, 35)</p> | $x_2 = \frac{18 + 19 + 21 + 23 + 25 + 26}{6} = 22$ $y_2 = \frac{24 + 20 + 17 + 14 + 11 + 4}{6} = 15$ <p style="text-align: center;">P₂(22, 15)</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $x_1 = \frac{7 + 9 + 11 + 14 + 15 + 16}{6} = 12$ $y_1 = \frac{45 + 40 + 36 + 34 + 30 + 25}{6} = 35$ <p style="text-align: center;">P₁(12, 35)</p> | $x_2 = \frac{18 + 19 + 21 + 23 + 25 + 26}{6} = 22$ $y_2 = \frac{24 + 20 + 17 + 14 + 11 + 4}{6} = 15$ <p style="text-align: center;">P₂(22, 15)</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4. La droite de régression est celle qui passe par les points P ₁ et P ₂ . | <p>Équation de la droite passant par les points P₁(12, 35) et P₂(22, 15) :</p> $y = ax + b$ $a = \frac{15 - 35}{22 - 12} \text{ et } y = -2x + b$ $= \frac{-20}{10} \quad 35 = -2 \times 12 + b$ $= -2 \quad b = 59$ <p>Équation de la droite de régression : $y = -2x + 59$</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

MÉTHODE DE LA DROITE MÉDIANE-MÉDIANE

| | | | | | | | | | | | | | | | |
|--|---|--------------------------|---------|----------|----------|--|--------------------------|---------|---------|--|--------------------------|----------|----------|---------|----|
| Démarche | Exemple: | | | | | | | | | | | | | | |
| | (12, 39) | (9, 28) | (8, 23) | (17, 49) | (21, 53) | (3, 10) | (14, 37) | (6, 15) | (9, 26) | (4, 13) | (19, 48) | (16, 44) | (11, 32) | (8, 19) | |
| 1. Ordonner les couples de la distribution selon leurs abscisses. | Distribution à deux variables | | | | | | | | | | | | | | |
| | x | 3 | 4 | 6 | 8 | 8 | 9 | 9 | 11 | 12 | 14 | 16 | 17 | 19 | 21 |
| y | 10 | 13 | 15 | 19 | 23 | 26 | 28 | 32 | 39 | 37 | 44 | 49 | 48 | 53 | |
| 2. Diviser l'ensemble des couples en trois groupes égaux. Si ce n'est pas possible, les diviser pour que le premier et le dernier groupe comptent le même nombre de couples. | x | 3 | 4 | 6 | 8 | 8 | 9 | 9 | 11 | 12 | 14 | 16 | 17 | 19 | 21 |
| | y | 10 | 13 | 15 | 19 | 23 | 26 | 28 | 32 | 39 | 37 | 44 | 49 | 48 | 53 |
| | | Groupe 1 de 5 couples | | | | | Groupe 2 de 4 couples | | | | Groupe 3 de 5 couples | | | | |
| 3. Déterminer l'abscisse médiane et l'ordonnée médiane de chaque groupe pour former les couples médians $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ et $M_3(x_3, y_3)$. Il importe d'ordonner les ordonnées dans chaque groupe pour en déterminer les médianes. | • Médiane des abscisses: $x_1 = 6$ | | | | | • Médiane des abscisses: $x_2 = \frac{9+11}{2} = 10$ | | | | • Médiane des abscisses: $x_3 = 17$ | | | | | |
| | • Médiane des ordonnées: $y_1 = 15$ | | | | | • Médiane des ordonnées: $y_2 = \frac{28+32}{2} = 30$ | | | | • Médiane des ordonnées: $y_3 = 48$ | | | | | |
| | $M_1(6, 15)$ | | | | | $M_2(10, 30)$ | | | | $M_3(17, 48)$ | | | | | |
| 4. Déterminer les coordonnées du point P correspondant respectivement à la moyenne des abscisses et à la moyenne des ordonnées des points M_1 , M_2 et M_3 . | $P\left(\frac{6+10+17}{3}, \frac{15+30+48}{3}\right) = P(11, 31)$ | | | | | | | | | | | | | | |
| 5. Déterminer la pente de la droite qui passe par les points M_1 et M_3 . | La pente de la droite qui passe par les points $M_1(6, 15)$ et $M_3(17, 48)$ est: $\frac{48-15}{17-6} = 3$ | | | | | | | | | | | | | | |
| 6. La droite de régression est celle qui passe par le point P et qui a la même pente que la droite passant par les points M_1 et M_3 . | Équation de la droite passant par $P(11, 31)$ et dont la pente est 3: $y = 3x + b$ $31 = 3 \times 11 + b$ $b = -2$ | | | | | | | | | | | | | | |
| | Équation de la droite de régression: $y = 3x - 2$ | | | | | | | | | | | | | | |
| |  | | | | | | | | | | | | | | |

RENFORCEMENT**7.5** Droite de régression

1 Dans chaque cas, déterminez l'équation de la droite de régression à l'aide de la méthode médiane-médiane.

a)

| | | | | | | | | | | | | |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|
| x | 17 | 18 | 20 | 21 | 22 | 22 | 24 | 26 | 28 | 29 | 31 | 32 |
| y | 35 | 31 | 31 | 30 | 31 | 31 | 29 | 23 | 23 | 22 | 18 | 18 |

b)

| | | | | | | | | | | |
|----------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| x | 15 | 20 | 23 | 27 | 28 | 28 | 39 | 42 | 45 | 47 |
| y | 96 | 108 | 110 | 122 | 134 | 138 | 144 | 152 | 158 | 170 |

2 Chacun des couples de la table de valeurs ci-dessous indique les années d'expérience d'un employé d'une usine et son salaire (en k\$).

Expérience et salaire d'un employé d'une usine

| | | | | | | | | | | | | |
|----------------------------|----|----|----|----|----|----|----|----|----|----|----|----|
| Années d'expérience | 3 | 4 | 6 | 7 | 7 | 9 | 11 | 11 | 12 | 14 | 16 | 17 |
| Salaire (k\$) | 31 | 36 | 34 | 37 | 34 | 39 | 41 | 42 | 39 | 41 | 43 | 44 |

Selon ces informations, quel est le salaire d'un employé ayant cumulé 23 années d'expérience ?

Réponse: _____

3 Dans chaque cas, déterminez l'équation de la droite de régression à l'aide de la méthode de la droite de Mayer.

a)

| | | | | | | | | | | | | | | |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| x | 15 | 17 | 19 | 20 | 22 | 23 | 24 | 29 | 30 | 31 | 34 | 36 | 38 | 40 |
| y | 16 | 19 | 22 | 28 | 31 | 34 | 39 | 44 | 48 | 52 | 56 | 59 | 61 | 65 |

b)

| | | | | | | | | | | | | |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|
| x | 6 | 9 | 12 | 15 | 21 | 27 | 33 | 36 | 42 | 44 | 46 | 51 |
| y | 82 | 76 | 70 | 72 | 58 | 50 | 42 | 38 | 30 | 22 | 16 | 8 |

4 Chacun des couples de la table de valeurs ci-dessous indique les investissements et les revenus (en M\$) d'une entreprise pour chacune des dix dernières années.

Investissements et revenus annuels d'une entreprise

| | | | | | | | | | | |
|------------------------------|----|----|----|----|----|----|----|----|----|----|
| Investissements (M\$) | 1 | 2 | 3 | 4 | 5 | 5 | 6 | 7 | 8 | 9 |
| Revenus (M\$) | 18 | 21 | 23 | 25 | 28 | 31 | 34 | 37 | 38 | 40 |

À l'aide de ces informations, quel montant l'entreprise devrait-elle investir pour espérer toucher des revenus de 51 M\$ la prochaine année ?

Réponse: _____

ENRICHISSEMENT**7.5** Droite de régression

- 1** Après avoir analysé la distribution statistique à deux variables représentée dans la table de valeurs ci-dessous, Léa affirme que l'équation de la droite de régression est la même si on utilise la méthode de la droite médiane-médiane ou la méthode de la droite de Mayer. Démontrez qu'elle a raison.

| | | | | | | | | | | | | |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|
| x | 3 | 5 | 5 | 8 | 10 | 14 | 14 | 15 | 16 | 17 | 17 | 20 |
| y | 34 | 32 | 28 | 27 | 25 | 22 | 23 | 21 | 20 | 19 | 17 | 14 |

Réponse: _____