Procédure Numworks
Statistiques à 2variables

Menu : Régression
X1 : Rang Y2 : données
Graphique
OK : pour choix de la régression
OK : pour équation , R², changement de régression ou prédiction sachant X ou Y

**Activité TP**

Durant la pandémie, l’évolution du virus a subi des variations très importantes. Malheureusement, le 1er décès a eu lieu le 15 février 2020, le 2ème le 26 février.
Nous voulons vérifier l’évolution mathématiques du nombre de décès au début de la pandémie et essayer de comprendre le rôle et les difficultés de certains experts durant cette crise. Pour cela nous allons travailler sur la série statistique à partir du 25 février 2020, veille du 2ème décès :

|  |  |  |
| --- | --- | --- |
|  | Février | mars |
| Date | 25 | 26 | 27 | 28 | 29 | 1er |  | 15 |
| Rang | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nombre de décès | 1 | 2 | 2 | 2 | 2 | 2 | 3 | 4 | 4 | 7 | 11 | 16 | 19 | 30 | 33 | 48 | 61 | 79 | 91 | 127 |

1. En utilisant les fonctionnalités de la calculatrice « Numworks », complétez le tableau suivant :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Types de Régression** | Linéaire | QuadratiquePolynomiale de degré 2 | Exponentielle | logarithmique |
| **Adéquation de la courbe par rapport au nuage de points** |  |  |  |  |
| **Coefficient de détermination R²** |  |  |  |  |

1. En analysant les courbes et le coefficient de détermination, Quelle est la régression la plus pertinente ?
Donnez l’équation et le R² de cette courbe de tendance.

La fonction *ex* est une fonction particulière des fonctions exponentielles. On a e1 = e $≈2,71828$
*e* est une valeur particulière en mathématiques comme π.
La fonction *ex* et la fonction *ln(x)* (logarithme népérien) sont des fonctions réciproques comme les fonctions *10x* et *log(x)* et ayantles mêmes propriétés que *10x* et *log(x)*.

1. D’après notre modèle, le 15 mars, quelles étaient les prévisions de décès pour les 16, 17, 18 et 19 mars ?

|  |  |
| --- | --- |
|  | Mars |
| Date | 16 | 17 | 18 | 19 |
| Rang | 21 | 22 | 23 | 24 |  |
| Nombre de décès | 148 | 175 | 244 | 372 |  |

Voici les données réelles du 16 au 19 mars 2020  :

1. Faites une critique des résultats donnés par notre modèle de prévision

1. Le 15 mars, quand pouvions nous prévoir les 10 000 décès ? Vérifiez par un calcul. (Comparer ensuite avec les données réelles.)

1. Le 15 mars, le gouvernement annonçait le confinement général de la France le 16 mars. Comprenez-vous cette décision ? Vérifier l’impact de cette décision en comparant prévisions et données réelles sur les semaines suivantes.

**FICHE SYNTHÈSE**

1. **Série statistique à deux variables**



Une série statistique qui possède deux caractères quantitatifs est **dite à deux variables**. Ces valeurs sont de la forme **(*x*i ; *y*i).**

Dans un repère orthonormal, à chaque couple (*x*i ; *y*i) est associé un point Mi de coordonnées (*x*i ; *y*i).

L’ensemble des points obtenus s’appelle un nuage de points.

Exemple : Dans l’activité les deux caractères étudiés sont le rang du jour (mesure de temps) et le nombre de décès. C’est donc une série statistique chronologique à deux variables.

1. **Ajustement d’un nuage de points**

Ajuster un nuage de points, c’est trouver la modélisation mathématiques qui s’approche au mieux du nuage de points. Cela revient donc à chercher une fonction *f*, telle que sa courbe représentative, d’équation *y*=*f*(*x*), passe au plus près de l’ensemble des points du nuage.

Voici 4 types d’ajustement :

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| Ajustement affineou linéaire | Ajustement paraboliqueou quadratique | Ajustement exponentiel | Ajustement logarithmique |
| *y*=a*x*+b | *y*=a*x*²+b*x*+c | y=a×*e*b*x* | *y*=a×*ln*(*x*)+b |

La fonction *ex* est une fonction particulière des fonctions exponentielles. On a e1 = e $≈2,71828$
*e* est une valeur particulière en mathématiques comme π.
La fonction *ex* et la fonction *ln(x)* (logarithme népérien) sont des fonctions réciproques comme les fonctions *10x* et *log(x)* et ayantles mêmes propriétés que *10x* et *log(x)*.

1. **Prévisions, coefficient de détermination**

L’équation de la courbe, ou la représentation graphique permettent de faire des prévisions. On extrapole quand on cherche une valeur au-delà de l’intervalle d’étude, on interpole quand on cherche une valeur dans l’intervalle d’étude.

Le coefficient de détermination R² est une valeur comprise entre 0 et 1 calculée par le logiciel. Plus cette valeur est proche de 1, plus l’ajustement et donc la prévision seront pertinents. On dit que la corrélation est forte.

Exemple : Dans l’activité nous avons réalisé quatre ajustements. L’ajustement exponentiel était l’un dont le coefficient de détermination R² était proche de 1. Nous avons pu faire ainsi des prévisions.

**FICHE D’EXERCICES**

**Problème 1**

Un gérant possède 2 magasins de vente de vêtements ; le premier dans une zone commerciale en périphérie de la ville, et le second dans l’hypercentre dans une rue piétonne.
Voici les Chiffres d’affaires de ces 2 magasins les 5 derniers trimestres :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Trimestre | 1 | 2 | 3 | 4 | 5 |
| CA Magasin ZC | 102 552 € | 104 325 € | 105 758 € | 106 263 € | 106 548 € |
| CA Magasin HC | 103 000 € | 104 802 | 105 856 | 106 604 | 107 184 |

1. Pour le 1er magasin, représenter graphiquement l’étude par un nuage de points.
2. Trouver le meilleur ajustement pour ce nuage et donner son équation. Justifier ce choix.
3. Faire une prévision du chiffre d’affaires pour les trimestre 6, 7 et 8.
4. Répondre aux mêmes questions pour le 2nd magasin.
5. Dans le cas du second magasin, quand le chiffre d’affaires dépassera-t-il les 110 000 € si cette évolution se poursuit ? Trouver la réponse en résolvant une équation puis en vérifiant avec les fonctionnalités de la calculatrice.

**Problème 2**

Voici, ci-dessous, un tableau donnant l’évolution du nombre de centenaires depuis 1960 :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Année | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 |
| Rang X | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| Nombre Y | 977 | 1122 | 1 545 | 3 760 | 8 063 | 14 504 | 20 944 |

1. Représenter graphiquement cette étude par un nuage de points.
2. Quels sont les 3 meilleurs ajustements pour ce nuage ? Justifier.
3. Dans chacun des cas donner l’équation de cet ajustement.
4. En analysant les courbes de ces 3 ajustements sur [0 ; 150] pour X et [-10 000 ; 100 000] pour Y, expliquez pourquoi on choisira plutôt l’ajustement exponentiel.
5. Dans ce cas, quand dépassera-t-on les 100 000 centenaires ? Trouver la réponse en résolvant une équation puis en vérifiant avec les fonctionnalités de la calculatrice.
6. Quelle conclusion un gouvernement pourrait-il faire avec ce résultat ? Quelle conséquence cela pourrait-il avoir ?

**Problème 3** **ventes croisées / cross-selling / ventes additionnelles**

Dans un magasin, un réagencement a permis de disposer les accessoires/bijoux en bout de caisse. Voici 2 études des chiffres d’affaires mensuels en € faites avant et après le réagencement :

Avant l’agencement :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| CA en € du magasin X | 53282 | 48525 | 63 721 | 49524 | 54525 | 64257 |
| CA en € acc/bijoux Y | 1421 | 1532 | 985 | 2632 | 1127 | 1834 |

Après l’agencement :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| CA en € du magasin X | 48722 | 54235 | 61257 | 46533 | 59782 | 65723 |
| CA en € acc/bijoux Y | 1357 | 1683 | 2213 | 1246 | 2089 | 2634 |

1. Représenter graphiquement l’étude avant l’agencement, par un nuage de points.
2. Un ajustement vous paraît-il acceptable ? Justifier la réponse.
3. Représenter graphiquement l’étude après l’agencement, par un nuage de points.
4. Quel est le meilleur ajustement de ce nuage ? Justifier la réponse.
5. Donner l’équation de cet ajustement.
6. Quel intérêt a eu ce réagencement en terme de chiffre d’affaires ?
7. Quel chiffre d’affaires doit faire le magasin si on veut un chiffre d’affaire du rayon accessoires/bijoux dépassant les 3000 € ? Trouver la réponse en résolvant une équation puis en vérifiant avec les fonctionnalités de la calculatrice.

**Problème 4**

Une entreprise de vente en ligne a lancé une nouvelle gamme de produits et souhaite prédire les ventes futures en fonction du temps afin d'optimiser sa stratégie de marketing. La société dispose de données historiques sur les ventes de ces produits au cours des derniers mois. Elle vous demande de faire une prévision des ventes.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mois | 1 | 2 | 3 | 4 | 5 | 6 |
| Vente en € | 4721 | 5254 | 6254 | 7928 | 9425 | 11334 |

1. Représenter graphiquement cette étude par un nuage de points sur votre calculatrice.
2. Trouver les 2 meilleurs ajustements pour ce nuage de points. Justifier votre choix.
3. Faire une prévision des ventes pour les mois 7, 8 et 9 dans chacun des cas.

Voici finalement les données réelles des mois 7, 8 et 9.

|  |  |  |  |
| --- | --- | --- | --- |
| Mois | 7 | 8 | 9 |
| Vente en € | 13522 | 16148 | 19429 |

1. Faire une critique des 2 modèles précédents.
2. A partir de maintenant, on choisira le modèle exponentiel. Rajouter les données réelles des mois 7,8 et 9 sur votre calculatrice.
3. Donner l’équation du modèle exponentiel. Les coefficients seront arrondis à 2 chiffres après la virgule.
4. Prévoir **par le graphique et par le calcul** le mois où les ventes dépasseront les 40 000 €.
5. Rédiger pour cette entreprise, un rapport de votre étude, donnant votre démarche de travail, les premières prévisions, les modifications éventuelles de votre stratégie et le moment où l’on dépassera l’objectif des 40 000 € de vente. Vous préciserez aussi l’intérêt pour l’entreprise de pouvoir prévoir ces futures ventes.

**Problème 5**

Une entreprise de fabrication de vêtements souhaite prévoir ses coûts de production pour lancer une nouvelle collection. Pour ce faire, elle doit modéliser la relation entre le nombre d'articles produits et le coût de production total. C’est l’objet de votre travail. Voici les données fournies par cette entreprise.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nombre d’articles produits | 2000 | 4000 | 6000 | 8000 |
| Coût de production en € | 11050 | 12062 | 12651 | 13068 |

1. Représenter graphiquement cette étude par un nuage de points sur votre calculatrice.
2. Trouver le meilleur ajustement pour ce nuage de points. Justifier votre choix.
3. Donner l’équation du modèle. Les coefficients seront arrondis à 2 chiffres après la virgule.
4. Cette semaine, l’entreprise compte produire 5600 articles. Estimer son de coût de production.
5. L’entreprise ne peut dépasser les 16 000 € de coût de production. Estimer par le calcul, le nombre maximal d’articles à produire.
6. Vérifier votre résultat à la calculatrice et faire une critique sur le résultat précédent.

**Problème 6**

Un hôpital souhaite analyser le temps d'attente moyen des patients aux urgences au fil du temps pour améliorer leur gestion des ressources et réduire les délais. Voici les données sur plusieurs années :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Années | 1 | 2 | 3 | 4 | 5 | 6 |
| Temps d’attente moyenen minutes | 41 | 52 | 65 | 81 | 101 | 126 |

1. Représenter graphiquement cette étude par un nuage de points sur votre calculatrice.
2. Trouver le meilleur ajustement pour ce nuage de points. Justifier votre choix.
3. Donner l’équation du modèle. Les coefficients seront arrondis à 2 chiffres après la virgule.
4. Si rien n’ai fait, prévoir le temps d’attente moyen des patients pour les années 7,8 et 9.
5. Quand le temps d’attente dépassera-t-il les 6h ? Vous répondrez à cette question en résolvant une équation.
6. Vérifier votre résultat à la calculatrice.
7. Rédiger un rapport à l’hôpital sur cette étude, exposant la méthode utilisée et les prévisions obtenues.
Vous conclurez aussi sur l’acceptabilité de ces futurs temps d’attente et proposerez des solutions éventuelles.