

Thème : Séries statistiques à deux variables

Cet énoncé est celui de la deuxième épreuve orale (épreuve sur dossier) du Capes Externe de mathématiques, proposé aux candidat(e)s le 10 Juillet 2005.

Pour consulter les archives de cette épreuve orale, depuis la session 2005, on se reportera au site officiel du jury, à l'adresse <http://capes-math.org/>

1. L'exercice proposé au candidat

Le tableau ci-dessous donne la production annuelle d'une usine de pâte à papier (en tonnes) en fonction de l'année :

Année	1996	1997	1998	1999	2000	2001	2002	2003
Production	325	351	382	432	478	538	708	930

1. Tracer le nuage de points correspondant.
2. Pour l'année i , on note p_i la production de pâte à papier et $\ell_i = \ln(p_i)$.
Tracer le nouveau nuage de points $(i; \ell_i)$.
3. En utilisant la calculatrice, donner une équation de la droite d'ajustement par les moindres carrés de ℓ_i en i .
4. En déduire une fonction d'ajustement de la production en fonction de l'année.
5. Quelle production peut-on prévoir en 2005 ?

2. Le travail demandé au candidat


En aucun cas, le candidat ne doit rédiger sur sa fiche sa solution de l'exercice. Celle-ci pourra néanmoins lui être demandée partiellement ou en totalité lors de l'entretien avec le jury

Après avoir résolu et analysé l'exercice le candidat rédigera sur sa fiche les réponses aux questions suivantes :

1. Indiquer les classes de Lycée dans lesquelles on peut proposer cet exercice et les notions et outils mis en œuvre dans sa résolution.
2. Présenter sur une calculatrice les deux nuages de points.
3. Comment utiliseriez vous cet exercice pour présenter à une classe de Terminale ES la méthode d'ajustement par les moindres carrés ?
4. Quelles indications ajouteriez-vous à la question 4. pour amener un élève de Terminale à la résoudre ?
5. Proposer un ou plusieurs exercices sur le même thème.

Proposition de corrigé avec le Classpad 300

I. L'exercice proposé au candidat

- On ouvre l'application . Si c'est nécessaire, on supprime les données existantes par « Edit/Tout effacer ». On obtient ainsi l'écran de départ (fig1).

Dans la cellule marquée `liste1`, on entre le nom `annee`.

Dans la cellule `Cal` de cette première colonne, on entre `seq(k,k,1996,2003)`.

On obtient alors le tableau représenté (fig2).

Dans la cellule marquée `liste2`, on entre par exemple `prodn` (le nom `prod` est réservé par le système, et `production` fait plus de huit caractères).

Dans les cellules de cette colonne, on entre une à une les données de l'énoncé (fig3).

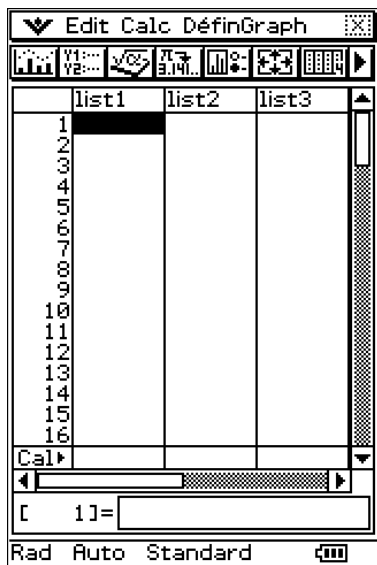


fig1 : écran stats initial

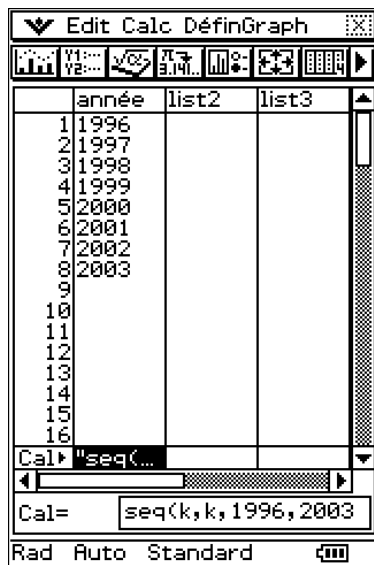


fig2 : colonne « année »

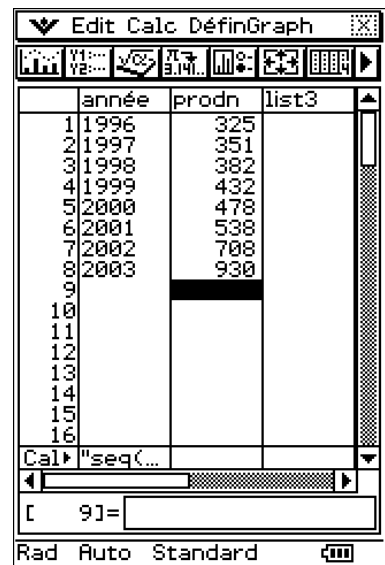




fig3 : colonne « prodn »


Avant de tracer le nuage, on choisit l'icône  (ou « DéfinGraph/Réglage ») (fig4), pour définir les caractéristiques du premier tracé statistique (sur neuf possibles).


Dans les listes déroulantes `listX` et `listY`, on choisit les listes `annee` et `prodn` (elles sont dans le répertoire courant, ici `main`). Les autres réglages par défaut de cette fenêtre nous conviennent (fig5).

On valide cette fenêtre, on choisit , et on obtient le nuage de points (fig6).

- On ferme la fenêtre de tracé par un appui sur .

Dans la cellule nommée `list3` on entre le nom `lnprod`.

Dans la cellule `Cal` de cette colonne, on entre l'expression `ln(prodn)` et on valide. Une action sur l'icône  force le passage en mode numérique approchée (on pourrait aussi passer en mode `Decimal` plutôt qu'en mode `Standard`). La troisième colonne contient maintenant les valeurs $\ell_i = \ln(p_i)$ (fig 7).

On va conserver le graphique n°1 pour mémoire (bien qu'il ne serve plus dans l'énoncé) et définir un graphique n°2. On choisit  (ou bien sûr « DéfinGraph/Réglage ») et on passe dans l'onglet n°2, que l'on complète comme indiqué (fig8).

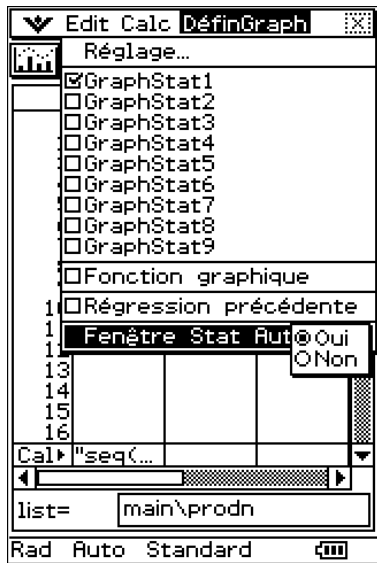


fig4 : menu DéfinGraph

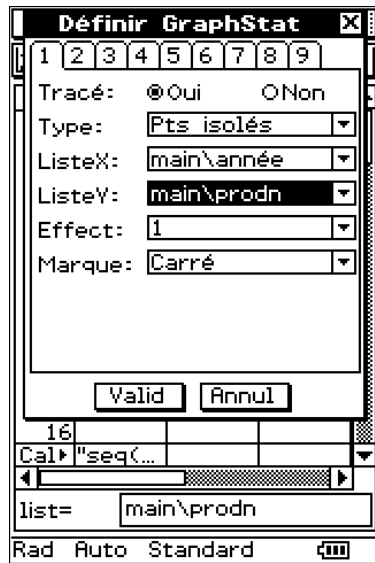


fig5 : Fenêtre Réglage

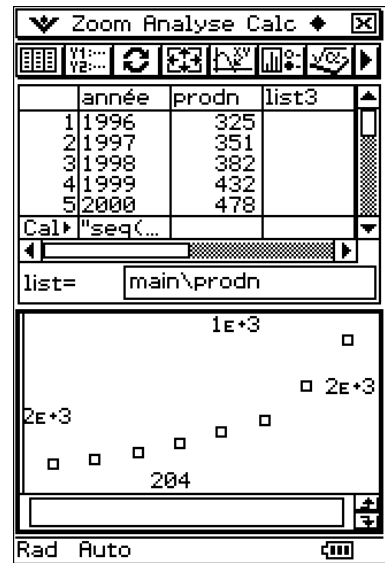



fig6 : le nuage de points

Dans le menu DéfinGraph, on vérifie que seul GraphStat2 est sélectionné.

On lance le tracé du nuage de points par un appui sur  (fig9).

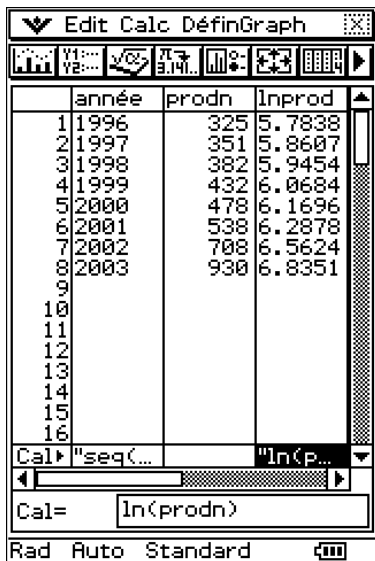


fig7 : colonne des $\ell_i = \ln(p_i)$

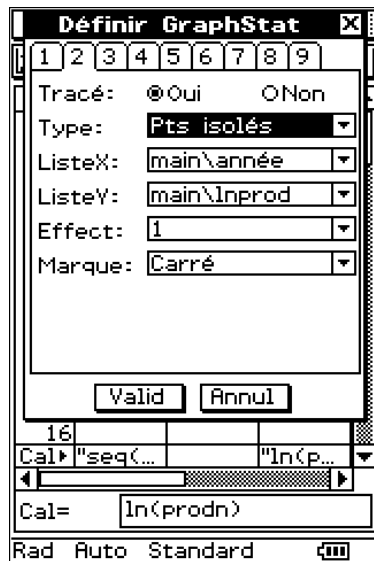


fig8 : déf^m graphique n°2

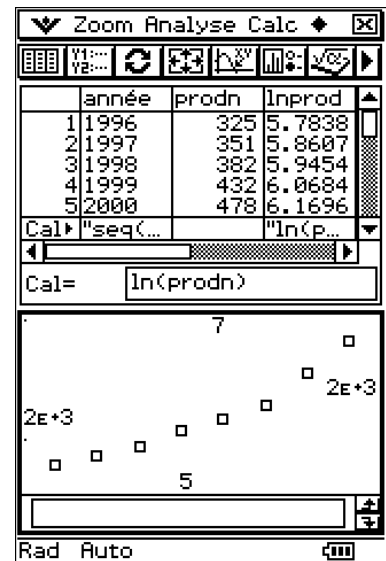



fig9 : nuage des (i, ℓ_i) .

3. On revient à l'affichage de la table des données.

On ouvre le menu Calc dans lequel on choisit « Rég linéaire » (fig10).

On complète la fenêtre de saisie comme indiqué (fig11).

On obtient alors l'équation de la droite d'ajustement par les moindres carrés de ℓ_i en i , sous la forme $y = ax + bx$, avec affichage des coefficients a, b . Le coefficient r , assez proche de 1, indique que l'approximation est de bonne qualité (fig12).

Immédiatement après un appui sur , la fenêtre graphique est affichée et on y voit un tracé conjoint du nuage des (i, ℓ_i) et la droite des moindres carrés (fig13).

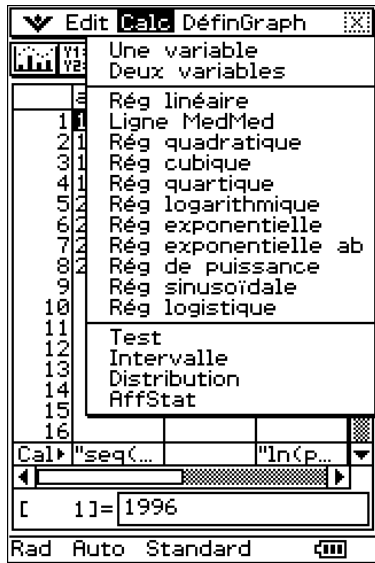


fig10 : le menu Calc

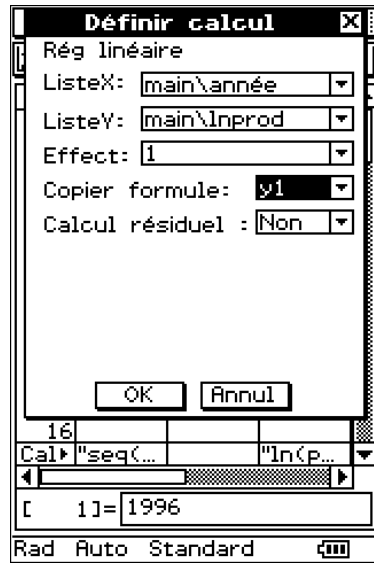


fig11 : choix des paramètres

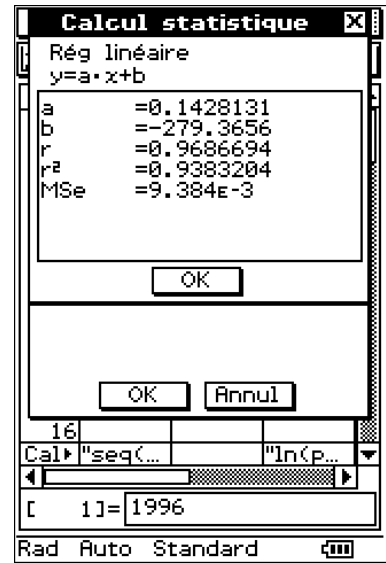


fig12 : moindres carrés

4. On a vu (fig11) qu'on avait demandé au Classpad de placer l'équation de la droite des moindres carrés dans la variable y1. Depuis l'écran graphique ou l'écran des données, un appui sur l'icône ouvre l'application principale.

On peut alors y afficher le contenu de la variable y1. Les coefficients a et b de l'équation $y = ax + b$ se trouvent par ailleurs dans les variables aCoef et bCoef (fig14).

Bien sûr on obtient ici une fonction d'ajustement de $\ln(p_i)$ en fonction de i .

On a $\ln(p_i) \approx ai + b$ donc $p_i \approx e^{ai+b}$ (fig14).

Remarque : le calcul a été effectué en mode Decimal et dans le format Fix 2 (pour ce genre de réglage, ouvrir le menu Format de base).

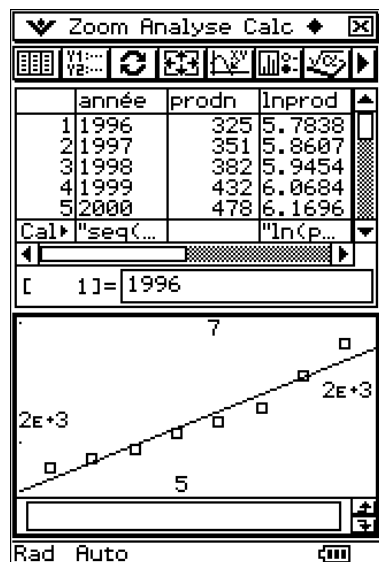


fig13 : droite et nuage

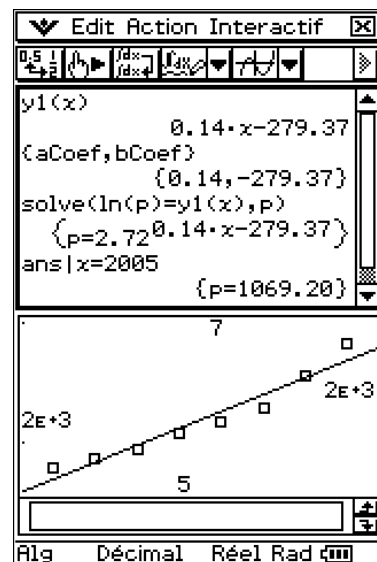


fig14 : en 2005 ?

5. Toujours dans l'application principale, il suffit d'évaluer (avec $x = 2005$) l'expression obtenue précédemment pour constater que la production en 2005 devrait (a dû, puisque ce corrigé est écrit en décembre 2007) atteindre 1069 tonnes.

II. Un peu de calcul formel avec le Classpad

Avec le calcul formel du Classpad, on va retrouver l'équation de la droite d'ajustement (au sens des moindres carrés) des $\ln(p_i)$ en fonction des années i .

Bien sûr, la justification de ce qui suit est un peu hors de portée des élèves de terminale (principalement en ce qui concerne l'utilisation des dérivées partielles) mais devrait être bien clair pour tout(e) candidat(e) au Capes.

On sait qu'il s'agit de trouver les réels a et b qui minimisent la somme $\sum_i (\ln(p_i) - ai - b)^2$, où l'entier i décrit l'ensemble des années de 1996 à 2003.

On ouvre donc l'application .


On sait que les variables `année` et `lnprod` sont déjà présentes en mémoire.

L'instruction `L:=lnprod-(a*année+b)` place dans la variable L la liste des $\ln(p_i) - (ai + b)$.

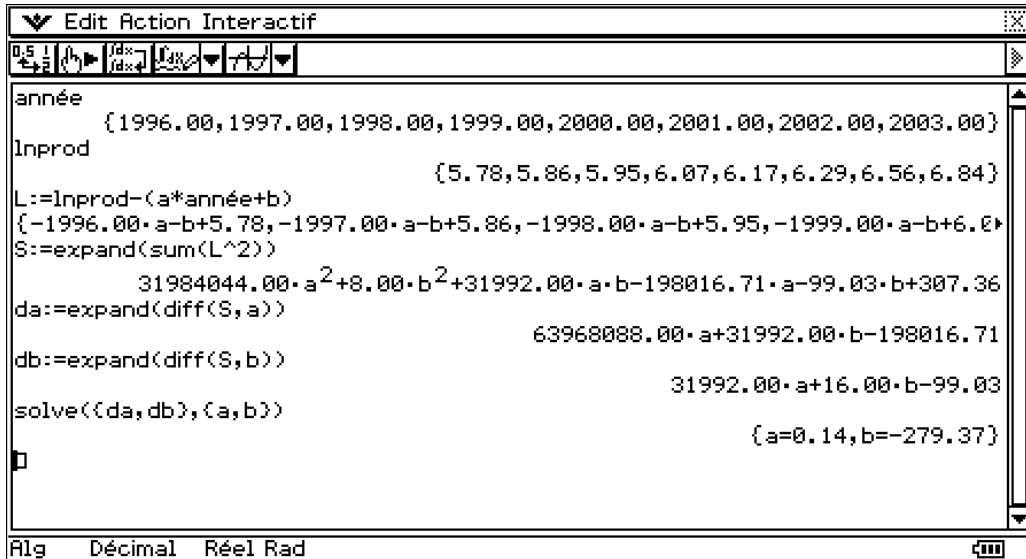
L'instruction `S:=expand(sum(L^2))` calcule (et développe) la somme des carrés des coefficients de la liste L . On vient donc de calculer $S = \sum_i (\ln(p_i) - ai - b)^2$.

Pour savoir pour quelles valeurs de a et de b cette expression est minimale, on calcule les dérivées partielles de S par rapport à a et à b .

L'instruction `solve` détermine ensuite le couple (a, b) qui annule ces deux dérivées.

On retrouve bien les coefficients a et b obtenus dans l'application  (fig15).

Remarque : pour plus de lisibilité, les calculs sont affichés dans l'écran du « Classpad Manager ».



```

Edit Action Interactif
-----
année      {1996.00,1997.00,1998.00,1999.00,2000.00,2001.00,2002.00,2003.00}
lnprod     {5.78,5.86,5.95,6.07,6.17,6.29,6.56,6.84}
L:=lnprod-(a*année+b)
{-1996.00*a-b+5.78,-1997.00*a-b+5.86,-1998.00*a-b+5.95,-1999.00*a-b+6.07}
S:=expand(sum(L^2))
31984044.00*a^2+8.00*b^2+31992.00*a*b-198016.71*a-99.03*b+307.36
da:=expand(diff(S,a))
63968088.00*a+31992.00*b-198016.71
db:=expand(diff(S,b))
31992.00*a+16.00*b-99.03
solve((da,db),(a,b))
{a=0.14,b=-279.37}
-----
Alg  Décimal  Réel Rad

```