Thème : Séries statistiques à deux variables

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 $l_i = \ln(p_i)$. Tracer le nouveau nuage de points (i, l_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 :

Q.1) Indiquer les classes de Lycée dans lesquelles on peut proposer cet exercice et les notions et outils mis en œuvre dans sa résolution.

Q.2) Présenter sur une calculatrice les deux nuages de points.

Q.3) Comment utiliseriez vous cet exercice pour présenter à une classe de Terminale ES la méthode d'ajustement par les moindres carrés ?

Q.4) Quelles indications ajouteriez-vous à la question 4. pour amener un élève de Terminale à la résoudre ?

Q.5) Proposer un ou plusieurs exercices sur le même thème.

3. Quelques références aux programmes

Programme de Terminale ES

Contenus	**Modalités de mise en œuvre**	**Commentaires**
Statistiques et probabilités Nuage de points associé à une série statistique à deux variables numériques. Point moyen.	On proposera aussi des exemples où la représentation directe en (x, y) n'est pas possible et où il convient, par exemple, de représenter $(x, \ln(y))$ ou $(\ln(x), y)$ et on fera le lien avec des repères semi-logarithmiques.	
Ajustement affine par des moindres carrés.	On fera percevoir le sens de l'expression "moindres carrés" par le calcul sur tableur, pour un exemple simple de la somme $(y_i - ax_i - b)^2$. On évoquera sur des exemples l'intérêt éventuel et l'effet d'une transformation affine des données sur les paramètres a et b. On étudiera avec des simulations la sensibilité des paramètres aux valeurs extrèmes. On proposera des exemples où une transformation des données conduit à proposer un ajustement affine sur les données transformées.	L'objectif est de faire des interpolations ou des extrapolations. On admettra les formules donnant les paramètres de la droite des moindres carrés : coefficient directeur et ordonnée à l'origine. On traitera essentiellement des cas où, pour une valeur de x, on observe une seule valeur de y (par exemple, les séries chronologiques). Le coefficient de corrélation linéaire est hors programme (son interprétation est délicate, notamment pour juger de la qualité d'un ajustement affine).
	On proposera un ou deux exemples où les points (x_i, y_i) du nuage sonr "presque" alignés et où cet alignement peut s'expliquer par la dépendance "presque" affine à une troisième variable.	On verra ainsi que pouvoir prédire y à partir de x ne prouve pas qu'il y ait un lien de causalité entre x et y.