BANQUE COMMUNE D'ÉPREUVES

Code sujet : 283

Conception : HEC Paris – ESCP BS

OPTION SCIENTIFIQUE

MATHÉMATIQUES II

Jeudi 29 avril 2021, de 8 h. à 12 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à ***encadrer*** *dans la mesure du possible les résultats de leurs calculs.*
Aucun document n'est autorisé. ***L'utilisation de toute calculatrice et de tout matériel électronique est interdite.*** *Seule l'utilisation d'une règle graduée est autorisée.*
Si au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

Partie 1 – Polynômes factoriels

On note F, l'espace vectoriel des fonctions polynomiales à coefficients réels et, pour tout entier naturel r, on note F_r, le sous-espace vectoriel des fonctions polynomiales à coefficients réels de degré inférieur ou égal à r.
On note U_k la fonction $x \mapsto x^k$ avec la convention habituelle $x^0 = 1$, de telle sorte que la base canonique de F_r est notée $(U_0, U_1, \ldots, U_r)$.

1) Soit r, un entier naturel. On considère une famille $(Q_0, Q_1, \ldots, Q_r)$ de fonctions polynomiales de degrés respectifs $d_0, d_1, \ldots, d_r$ avec $d_0 < d_1 < \cdots < d_r$.

1.a) On suppose qu'il existe des réels $\lambda_0, \lambda_1, \ldots, \lambda_r$ non tous nuls tels que

$$\lambda_0 Q_0 + \lambda_1 Q_1 + \cdots + \lambda_r Q_r = 0.$$

En considérant $m = \max\{k \in [\![0, r]\!] \ : \ \lambda_k \neq 0\}$, démontrer que l'hypothèse précédente est absurde.
Qu'a-t-on ainsi démontré ?

1.b) À quelle condition la famille $(Q_0, Q_1, \ldots, Q_r)$ est-elle une base de F_r ? (On précisera s'il s'agit d'une condition *nécessaire*, d'une condition *suffisante* ou d'une condition *nécessaire et suffisante*.)

2) Pour tout entier naturel r, le réel "x puissance r descendante" est noté $x^{\underline{r}}$ et défini par

$$\forall\, x \in \mathbf{R}, \qquad x^{\underline{r}} = x(x-1) \times \cdots \times (x - r + 1) = \prod_{k=0}^{r-1} (x - k)$$

avec la convention $x^{\underline{0}} = 1$. On pose alors

$$\forall\, r \in \mathbf{N}, \qquad V_r \ : \ x \mapsto x^{\underline{r}}.$$

Il est clair que V_r appartient à F.

2.a) Quelles sont les racines de V_r ?

2.b) Démontrer que la famille $(V_0, V_1, \ldots, V_r)$ est une base de F_r.

Tournez la page S.V.P.

2. c) Démontrer que, pour tout entier $r \geqslant 2$,

$$x^{\underline{r}} \underset{x \to +\infty}{=} x^r - \frac{r(r-1)}{2} x^{r-1} + o(x^{r-1}).$$

(On pourra raisonner par récurrence sur r.)

3) Ici, l'entier $r \geqslant 1$ est fixé et on compare la famille $(V_k)_{0 \leqslant k \leqslant r}$ à la base canonique $(U_k)_{0 \leqslant k \leqslant r}$ de l'espace F_r des fonctions polynomiales de degré inférieur ou égal à r.

3. a) Démontrer qu'il existe une unique famille $\left(\sigma(r,k)\right)_{0 \leqslant k \leqslant r}$ de nombres réels tels que

$$U_r = \sum_{k=0}^{r} \sigma(r,k) V_k.$$

3.b) Établir les relations suivantes :

$$\begin{aligned}
\forall\, r \in \mathbf{N}^*, \qquad & \sigma(r,0) = 0 && (1)\\
& \sigma(r,1) = \sigma(r,r) = 1 && (2)\\
& \sigma(r,r-1) = \frac{r(r-1)}{2} && (3)\\
\forall\, r \in [\![2,+\infty[\![, \qquad & \sigma(r,2) = 2^{r-1} - 1 && (4)
\end{aligned}$$

3. c) Démontrer que, pour tout $r \in \mathbf{N}^*$ et pour tout $k \in [\![1,r]\!]$, on a

$$\sigma(r+1,k) = \sigma(r,k-1) + k\,\sigma(r,k).$$

3.d) En déduire que $\sigma(r,k)$ est un entier naturel non nul pour tout $r \in \mathbf{N}^*$ et pour tout $k \in [\![1,r]\!]$.

3. e) Écrire un code scilab qui affiche (au moyen de la commande `disp`) les listes $\left(\sigma(r,k)\right)_{1 \leqslant k \leqslant r}$ pour r variant de 2 à 5.

On pourra utiliser la commande `ones(n,p)` qui retourne la matrice de $\mathfrak{M}_{n,p}(\mathbf{R})$ dont tous les coefficients sont égaux à 1.

4. a) Démontrer que, pour tout entier naturel r, il existe une unique famille $\left(s(r,k)\right)_{0 \leqslant k \leqslant r}$ de nombres réels tels que

$$\forall\, x \in \mathbf{R}, \qquad x^{\underline{r}} = \sum_{k=0}^{r} s(r,k) x^k.$$

4.b) Soit $r \in \mathbf{N}^*$. Démontrer que

$$\forall\, k \in [\![1,r]\!], \qquad s(r+1,k) = s(r,k-1) - r\,s(r,k).$$

En déduire la valeur de $s(r,1)$.

4. c) Déduire de **4.b)** que, pour tout $r \in \mathbf{N}^*$ et pour tout $k \in [\![1,r]\!]$, le signe de $s(r,k)$ est celui de $(-1)^{r+k}$.

4.d) Démontrer que $\sigma(r,r)s(r,r) = 1$ et que

$$\forall\, \ell \in [\![0,r-1]\!], \qquad \sum_{k=\ell}^{r} \sigma(r,k) s(k,\ell) = 0.$$

4. e) Calculer $s(r,r)$ pour tout $r \in \mathbf{N}^*$ et $s(r,r-1)$ pour tout entier $r \geqslant 2$.

Partie 2 – Quelques propriétés de la loi de Poisson

Sous réserve d'existence, on note respectivement $\mathbf{E}(A)$ *et* $\mathbf{V}(A)$*, l'espérance et la variance d'une variable aléatoire* A *et* $\mathbf{Cov}(A,B)$*, la covariance de deux variables aléatoires discrètes* A *et* B*.*

Dans cette partie, on note X*, une variable aléatoire définie sur un espace probabilisé* $(\Omega, \mathcal{A}, \mathbf{P})$ *qui suit la loi de Poisson de paramètre* $\theta > 0$*.*

Pour tout entier $r \geqslant 1$*, on pose* $X^{\underline{r}} = X(X-1) \times \cdots \times (X-r+1)$ *avec la convention* $X^{\underline{0}} = X^0 = 1$*.*

Avec la suite $\left(\sigma(r,k)\right)_{0 \leqslant k \leqslant r}$ *définie au* ***3.a)*** *et la suite* $\left(s(r,k)\right)_{0 \leqslant k \leqslant r}$ *définie au* ***4.a)****, on a*

$$\forall\, r \in [\![1,+\infty[\![, \qquad X^r = \sum_{k=0}^{r} \sigma(r,k) X^{\underline{k}} \quad \text{et} \quad X^{\underline{r}} = \sum_{k=0}^{r} s(r,k) X^k.$$

On admet ces deux résultats sans démonstration.

5.a) Rappeler sans démonstration les valeurs respectives de $\mathbf{E}(X)$, $\mathbf{V}(X)$ et $\mathbf{E}(X^2)$.

5.b) Exprimer X, X^2, X^3 et X^4 en fonction des variables aléatoires $X^{\underline{1}}$, $X^{\underline{2}}$, $X^{\underline{3}}$ et $X^{\underline{4}}$.

5.c) Démontrer que la variable aléatoire X admet des moments de tous ordres.

6.a) Justifier que, pour tout entier $r \geqslant 1$, la variable aléatoire $X^{\underline{r}}$ admet des moments de tous ordres.

6.b) Pour tout entier $r \geqslant 1$, exprimer $\mathbf{E}(X^{\underline{r}})$ en fonction de r et de θ.

6.c) Calculer $\mathbf{E}(X^3)$ et $\mathbf{E}(X^4)$ en fonction de θ.

7) Pour tout entier naturel k et pour tout réel $\theta > 0$, on pose

$$f(\theta, k) = \frac{\theta^k}{k!}\,\mathrm{e}^{-\theta} \qquad \text{et} \qquad g(\theta, k) = \ln\big(f(\theta, k)\big).$$

7.a) Pour tout entier $k \geqslant 0$, calculer l'expression de la dérivée partielle $\partial_1(g)(\theta, k)$ et exprimer la variable aléatoire $\partial_1(g)(\theta, X)$ en fonction de X et de θ.

7.b) Vérifier que $XX^{\underline{r}} = X^{\underline{r+1}} + rX^{\underline{r}}$ pour tout entier $r \geqslant 1$. En déduire que $\mathbf{Cov}(X, X^{\underline{r}}) = r\theta^r$.

7.c) Calculer $\mathbf{Cov}(\partial_1(g)(\theta, X),\ X^{\underline{r}})$ et en déduire l'inégalité

$$\forall\, \theta > 0,\ \forall\, r \in \mathbf{N}^*, \qquad \mathbf{V}(X^{\underline{r}}) \geqslant r^2\theta^{2r-1}.$$

Partie 3 – Estimation ponctuelle de fonctions du paramètre θ

Le contexte et les notations sont ceux de la partie 2.

On suppose que le paramètre $\theta \in \mathbf{R}_+^*$ est inconnu et on cherche ici à estimer $\varphi(\theta)$, où φ est une fonction dérivable sur $\mathbf{R}_+^*$.

Pour n entier de $\mathbf{N}^*$, on considère dans toute cette partie un n-échantillon $(X_1, \ldots, X_n)$ de variables aléatoires mutuellement indépendantes qui, comme X, suivent toutes la loi de Poisson de paramètre θ.

Pour tout $\theta > 0$ et pour tout $(k_1, \ldots, k_n) \in \mathbf{N}^n$, on pose

$$F(\theta, k_1, \ldots, k_n) = \mathbf{P}\Big(\bigcap_{i=1}^{n}[X_i = k_i]\Big) = \prod_{i=1}^{n} f(\theta, k_i)$$

et

$$G(\theta, k_1, \ldots, k_n) = \ln\big(F(\theta, k_1, \ldots, k_n)\big).$$

8.a) Démontrer que les fonctions $\theta \mapsto F(\theta, k_1, \ldots, k_n)$ et $\theta \mapsto G(\theta, k_1, \ldots, k_n)$ sont dérivables sur $\mathbf{R}_+^*$ et calculer les dérivées partielles $\partial_1(F)(\theta, k_1, \ldots, k_n)$ et $\partial_1(G)(\theta, k_1, \ldots, k_n)$.

8.b) Pour tout $\theta > 0$, on pose $Z_\theta = \partial_1(G)(\theta, X_1, \ldots, X_n)$. Démontrer que la variable aléatoire Z_θ est centrée et admet une variance strictement positive, notée $I(\theta)$, que l'on calculera.

On rappelle que : *s'il existe n séries absolument convergentes $\sum_{k_1 \in \mathbf{N}} v_{1,k_1}, \ldots, \sum_{k_n \in \mathbf{N}} v_{n,k_n}$ telles que*

$$\forall\, (k_1, \ldots, k_n) \in \mathbf{N}^n, \quad |u_{k_1,\ldots,k_n}| \leqslant |v_{1,k_1}| \times \cdots \times |v_{n,k_n}|$$

alors la série $\sum u_{k_1,\ldots,k_n}$ est dite <u>absolument convergente</u>. On admet que, dans ce cas, la somme

$$\sum_{(k_1,\ldots,k_n)\in\mathbf{N}^n} u_{k_1,\ldots,k_n}$$

est bien définie.

On rappelle l'énoncé de la Formule de transfert : *si la série* $\sum u_\theta(k_1,\ldots,k_n)F(\theta,k_1,\ldots,k_n)$ *est absolument convergente (au sens qui vient d'être rappelé), alors la variable aléatoire discrète* $U_\theta = u_\theta(X_1,\ldots,X_n)$ *est d'espérance finie et*

$$\begin{aligned}\mathbf{E}(U_\theta) &= \sum_{(k_1,\ldots,k_n)\in\mathbf{N}^n} u_\theta(k_1,\ldots,k_n)F(\theta,k_1,\ldots,k_n)\\ &= \sum_{(k_1,\ldots,k_n)\in\mathbf{N}^n} u_\theta(k_1,\ldots,k_n)\mathbf{P}\big([X_1=k_1]\cap\cdots\cap[X_n=k_n]\big).\end{aligned}$$

Soit $t : \mathbf{N}^n \to \mathbf{R}$*, une application indépendante de* θ*. On peut alors considérer la variable aléatoire discrète*

$$T = t(X_1,\ldots,X_n)$$

comme un estimateur de $\varphi(\theta)$*.*

On dira que la variable aléatoire T *est un* <u>*estimateur régulier*</u> *de* $\varphi(\theta)$ *lorsque les trois conditions suivantes* (R_1)*,* (R_2) *et* (R_3) *sont satisfaites.*

$$\mathbf{E}(T) = \varphi(\theta) \qquad (R_1)$$

$$\mathbf{V}(T) \text{ existe} \qquad (R_2)$$

$$\forall\,\theta>0,\quad \varphi'(\theta) = \sum_{(k_1,\ldots,k_n)\in\mathbf{N}^n} t(k_1,\ldots,k_n)\partial_1(F)(\theta,k_1,\ldots,k_n) \qquad (R_3)$$

On notera que la condition (R_3) *sous-entend que le second membre est la somme d'une série absolument convergente (au sens rappelé plus haut).*

9) Dans cette question, on suppose que T est un estimateur régulier de $\varphi(\theta)$.

9.a) La variable aléatoire T est-elle un estimateur sans biais de $\varphi(\theta)$?

9.b) Pourquoi la condition (R_3) n'est-elle pas une conséquence directe de la condition (R_1) ?

10) Soit T, un estimateur régulier du paramètre $\varphi(\theta)$.

10.a) Établir les égalités suivantes :

$$\forall\,\theta>0,\qquad \varphi'(\theta) = \mathbf{E}(T\times Z_\theta) = \mathbf{Cov}(T, Z_\theta).$$

10.b) En déduire l'inégalité

$$\forall\,\theta>0,\qquad \mathbf{V}(T) \geqslant \frac{\left(\varphi'(\theta)\right)^2}{I(\theta)}$$

où $I(\theta)$ a été défini au **8.b)**.

11) On cherche à simuler un échantillon de N réalisations de Z_θ pour différents couples (n,θ).

11.a) Compléter le code scilab suivant en justifiant votre réponse.

```
function ech=Z_th(n, theta)
    X=grand(n,N,'poi',theta);
    ech = (sum(X,      ) - n*theta)/theta
endfunction
```

On rappelle l'usage de la commande `sum` : pour un tableau $M = (M_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p}$, les deux instructions `sum(M, 'r')` et `sum(M, 'c')` retournent respectivement les tableaux

$$\left(\sum_{i=1}^{n} M_{i,j}\right)_{1\leqslant j\leqslant p} \qquad \text{et} \qquad \left(\sum_{j=1}^{p} M_{i,j}\right)_{1\leqslant i\leqslant n}$$

de tailles (*size*) respectives $(1,p)$ et $(n,1)$.

11.b) À l'aide de la commande `histplot`, on a tracé les histogrammes des échantillons obtenus pour les couples $(n, \theta) = (10, 4), (20, 4), (40, 4)$ et $(50, 5)$.
À quels couples correspondent les figures suivantes ? (On pourra admettre que $I(\theta) = n/\theta$.)

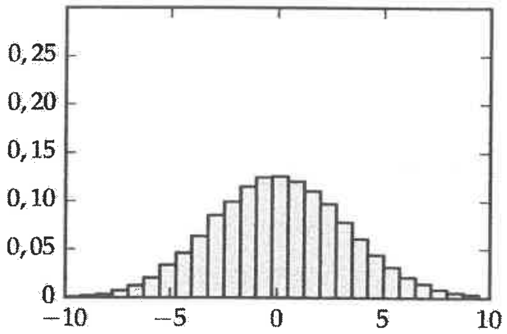

Figure A

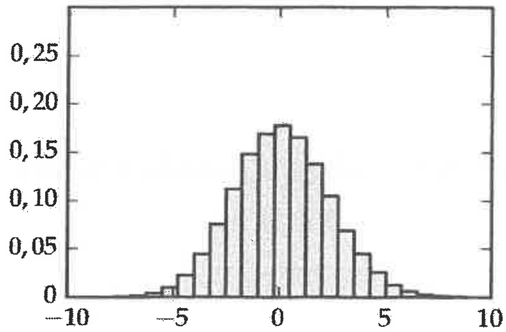

Figure B

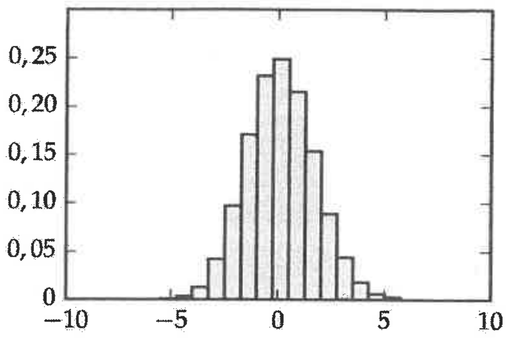

Figure C

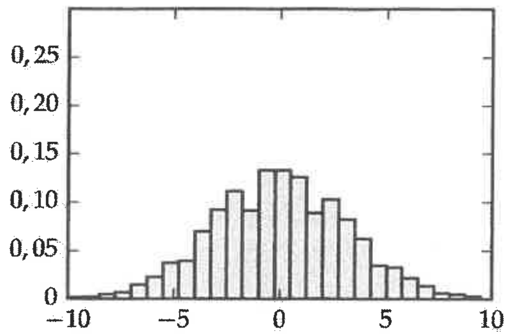

Figure D

12) Soit un entier $r \geqslant 1$. On suppose ici que $\varphi(\theta) = \theta^r$ et pour tout entier n supérieur ou égal à 1, on pose

$$S_n = X_1 + \cdots + X_n \quad \text{et} \quad M_{r,n} = \frac{S_n(S_n - 1) \times \cdots \times (S_n - r + 1)}{n^r}.$$

12.a) Rappeler (sans démonstration) la loi de la variable aléatoire S_n ainsi que son espérance et sa variance.

12.b) Démontrer que $M_{r,n}$ est un estimateur régulier de θ^r.
NB : Pour établir la propriété (R_3), on *admettra* que la série est absolument convergente.
En déduire que

$$\forall\, \theta > 0, \qquad \mathbf{V}(M_{r,n}) \geqslant \frac{r^2\theta^{2r-1}}{n}.$$

12.c) Dans cette question, on suppose que $r = 2$. Calculer la variance de $M_{2,n}$ et démontrer que la suite d'estimateurs de θ^2

$$(M_{2,n})_{n \geqslant 1}$$

est convergente.

12.d) Pour un entier $r \geqslant 1$ quelconque, la suite

$$(M_{r,n})_{n \geqslant 1}$$

d'estimateurs de θ^r est-elle convergente ? (On pourra commencer par calculer, en fonction de l'entier $k \in \mathbf{N}^*$ un équivalent de $\mathbf{E}(S_n^k)$ lorsque n tend vers $+\infty$.)

Tournez la page S.V.P.

Partie 4 – Le cas $\varphi(\theta) = \theta$

Le contexte et les notations sont ceux des parties 2 et 3.
Dans cette partie, on compare deux estimateurs du paramètre inconnu θ.

13) Pour tout entier n supérieur ou égal à 1, on pose

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

13.a) Démontrer que $\overline{X}_n$ est un estimateur régulier du paramètre θ.

13.b) Que devient l'inégalité du **10.b)** ?

On dit qu'un estimateur régulier de θ est <u>efficace</u> lorsque sa variance est minimale parmi les estimateurs réguliers de θ.

14) Soit Y, un estimateur régulier de θ. Pour tout réel α, on pose

$$\psi(\alpha) = \overline{X}_n + \alpha(Y - \overline{X}_n).$$

14.a) Vérifier que $\psi(\alpha)$ est un estimateur régulier de θ pour tout $\alpha \in \mathbf{R}$.

14.b) En déduire que

$$\mathbf{Cov}(\overline{X}_n, Y) = \frac{\theta}{n}.$$

14.c) Exprimer $\mathbf{V}(Y - \overline{X}_n)$ en fonction de $\mathbf{V}(Y)$ et de $\mathbf{V}(\overline{X}_n)$. En déduire qu'un estimateur efficace de θ est presque sûrement unique.

15) Pour tout entier $n \geqslant 2$, on pose

$$W_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$

15.a) Exprimer W_n en fonction de $\displaystyle\sum_{i=1}^{n} X_i^2$ et de $\overline{X}_n^2$.

15.b) Démontrer que W_n est un estimateur sans biais de θ.

15.c) Démontrer que W_n admet une variance (qu'on ne cherchera pas à calculer).

15.d) Étudier la convergence des deux suites d'estimateurs $(\overline{X}_n)_{n\geqslant 1}$ et $(W_n)_{n\geqslant 2}$ du paramètre inconnu θ.
On pourra démontrer que : si une suite réelle $(a_n)_{n\in\mathbf{N}}$ converge vers $a \in \mathbf{R}$ et si deux suites $(Y_n)_{n\in\mathbf{N}}$ et $(Z_n)_{n\in\mathbf{N}}$ de variables aléatoires convergent en probabilité vers les réels y et z respectivement, alors la suite de variables aléatoires $\big(a_n(Y_n - Z_n)\big)_{n\in\mathbf{N}}$ converge en probabilité vers le réel $a(y-z)$.

16) On simule des échantillons de N réalisations des estimateurs $\overline{X}_n$, W_n et

$$W_n' = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

avec $n = 50$. En comparant les figures suivantes, relier chaque histogramme à l'estimateur qui lui correspond.

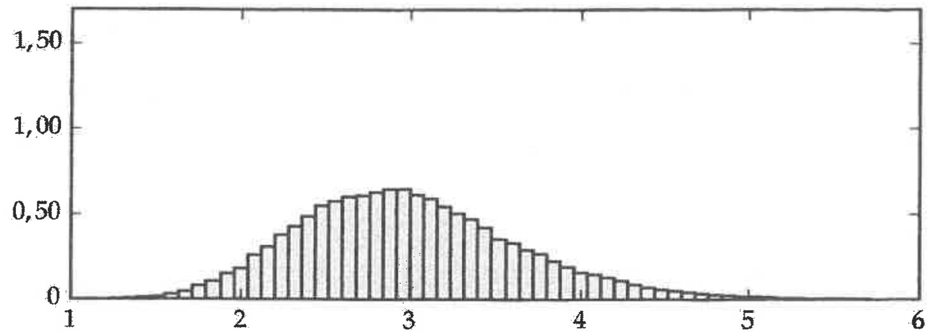

Figure E

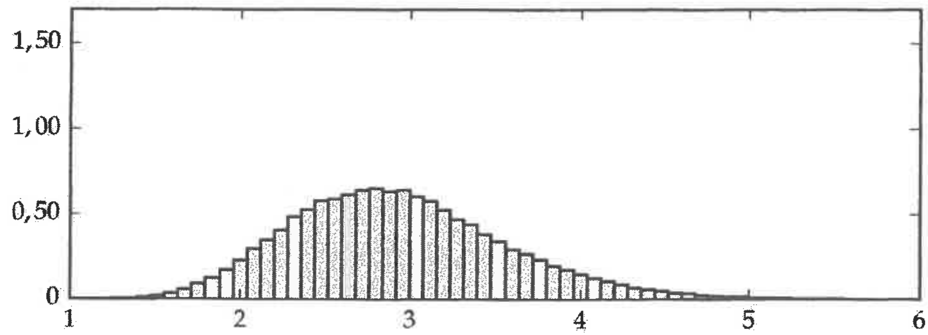

Figure F

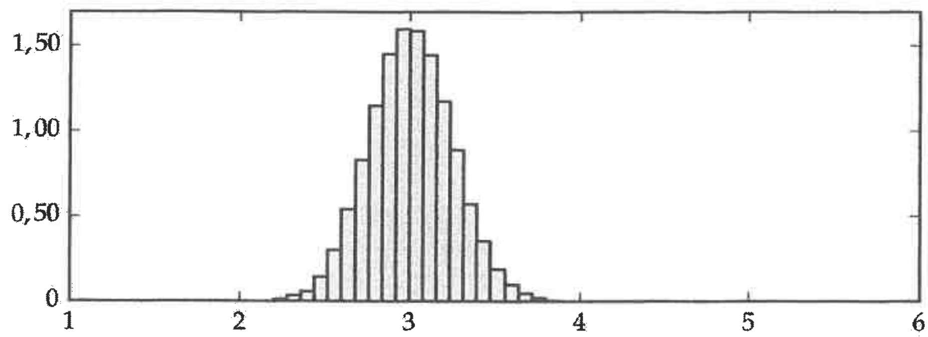

Figure G

Fin de l'énoncé