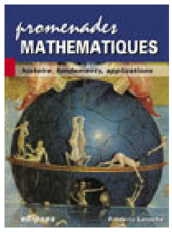


[accueil](http://)
Fichier Pdf

Approcher la loi Normale

La question est abordée sous forme de problème. La solution a été préparée pour le livre
Antisèches Mathématiques *de* **Terminale S**
mais est généreusement fournie ici.

Pierre Simon de Laplace fut un des premiers à utiliser les probabilités continues : on lui doit de nombreux résultats mathématiques dont certains obtenus **expérimentalement** !

Une question d'orthographe : depuis que j'enseigne j'ai toujours eu un doute sur l'orthographe de bino(ô)mial(e) ; quelques recherches succintes m'ont permis de voir que ce n'était pas si simple :

Depuis la dernière réforme de l'orthographe il n'y a pas de règle absolue pour la disparition ou non de l'accent circonflexe entre un mot et le substantif fabriqué à partir de lui :

tôle, tôlier ; pôle, polaire ; cône, conique ; soûler, soûlard…

Voici quelques orthographes pêchées au hasard :

binome : Appell, 1900 (Poincaré à la même époque n'utilise pas le terme) ; Chassé, Pavé, 1975. L'*usage* dans les manuels est d'utiliser *binôme*.

binômial : Angot, 1972, l'*usage* est d'utiliser *binomial*.

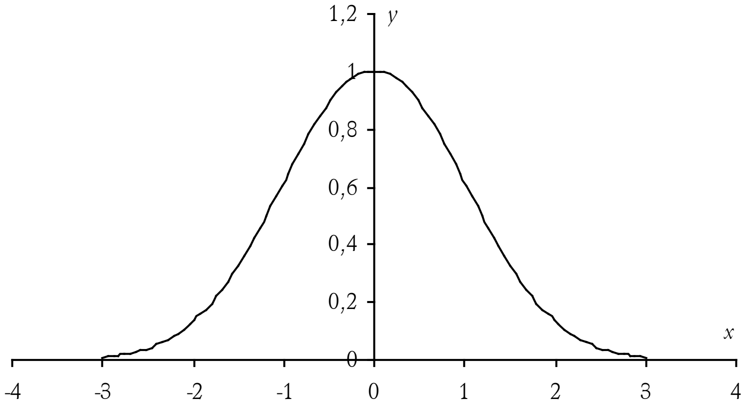
Le mot binomial (ou binômial) **n'existe pas** dans le dictionnaire de l'Académie Française, non plus que dans le Grand Robert, non plus que dans le Trésor de la Langue Française. On le trouve par contre dans le Larousse, sans accent.

B. Hauchecorne à propos de *binôme* et autres « nômes » fait la remarque qu'ils devraient en toute rigueur s'écrire *-*nome*. (*Les mots et les Maths, ed Ellipses*)

D'ailleurs l'Académie Française recommande la suppression dudit accent lorsque sa signification n'est pas claire (particulièrement avec *i* et *u*), ce qui semble bien le cas ici. Dorénavant je ne mettrai plus d'accent…

http://www.academie-francaise.fr/dictionnaire/

1. En utilisant la méthode d'Euler sur l'équation différentielle $y' + xy = 0$ (**1**) avec $y(0) = 1$ on a obtenu la courbe suivante sur $[-3 ; 3]$:



a. Soit $h>0$, montrer que l'équation (**1**) donne la relation $y(x+h)=y(x)(1-hx)$.

b. En partant de $x = 0$, montrer que cette relation devient $y(nh)=y(0)(1-hx)^n=(1-hx)^n$.

c. On pose $h=\dfrac{x}{n}$; donner alors l'expression de $y(x)$ en fonction de x et n.

d. En utilisant la définition de la fonction exponentielle, montrer que $y(x)=e^{-x^2}$. Montrer que cette fonction est paire ; qu'en déduit-on pour la résolution de (**1**) lorsque $h < 0$?

e. Etudier la fonction $f : x \to e^{-x^2}$ sur $\mathbb{R}$ et dresser son tableau de variations.

On voit sur la figure précédente que la fonction change de courbure en deux endroits (entre -2 et -1 puis entre 1 et 2) ; ceci signifie que pour certaine valeurs elle est au-dessus de sa tangente et pour d'autres elle est en dessous. Nous allons chercher les points où cela se produit d'une manière générale puis dans ce cas particulier.

2. Soit f une fonction quelconque définie et dérivable au moins deux fois en un point a ; g la fonction représentant sa tangente en a.

Considérons la fonction $u(x) = f(x) - g(x)$ définie par $u(x)=f(x)-f'(a)(x-a)-f(a)$.

a. Vérifier que $u'(x)=f'(x)-f'(a)$.

b. On suppose que $f''(x)>0$ sur un intervalle $I = [x_1 ; x_2]$ contenant a ; en déduire le tableau de variation de u' puis celui de u sur I (on pensera à calculer $u'(a)$ et $u(a)$…). Déterminer alors le signe de u.

c. Recommencer la même manœuvre lorsque $f''(x)<0$.

d. En déduire un théorème indiquant les positions relatives de la courbe de f et de sa tangente en a.

e. Utiliser ce théorème pour préciser les points où la fonction f du 1. change de courbure (on appelle ces points des *points d'inflexion*). Tracer les tangentes à la courbe de f en ces points puis tracer la courbe de f.

f. On modifie la fonction f en effectuant une translation et une homothétie sur les x : on considère alors

$$F(x)=e^{-\left(\frac{x-a}{b}\right)^2}$$

où a et b sont deux constantes. Préciser alors les trois points caractéristiques de la courbe de F : maximum et points d'inflexion. Faire quelques tracés avec diverses valeurs de a et b. Comment doit-on choisir b si l'on veut que la distance entre les points d'inflexion soit égale à 1 ?

3. On considère maintenant la suite $v_k = \binom{n}{k}(1-p)^k p^{n-k}$ où p est un réel compris entre 0 et 1, n un entier fixé et k un entier variant entre 0 et n.

a. A l'aide de votre calculatrice ou d'un tableur représenter les termes de la suite v_k dans les cas suivants :

$n = 5, p = 0{,}5$ $\quad$ $n = 5, p = 0{,}2$

$n = 20, p = 0{,}5$ $\quad$ $n = 20, p = 0{,}2$

$n = 100, p = 0{,}5$ $\quad$ $n = 100, p = 0{,}2.$

Que pensez-vous des représentations obtenues ? Démontrez que $\sum_{k=0}^{n} v_k = 1$.

b. On note $1 - p = q$; montrez que $\dfrac{\binom{n}{k+1}}{\binom{n}{k}} = \dfrac{n-k}{k+1}$ puis que $v_{k+1} = v_k \dfrac{(n-k)}{(k+1)}\dfrac{p}{q}$.

En déduire que $v_{k+1} > v_k$ lorsque $k < np + 1 - p$. Notons $k_0 =< np >$ l'entier le plus proche de np ; que peut-on dire de v_{k_0} ?

c. On va chercher une « dérivée » de la suite v_k en calculant le taux d'accroissement $\dfrac{v_{k+1} - v_k}{v_k}$: montrez que ce taux vaut $\dfrac{np - k - q}{(k+1)q}$.

d. On pose $x_k = \dfrac{k - np}{\sqrt{npq}}$ et $y_k = v_k\sqrt{npq}$; vérifier que $x_{k+1} - x_k = \dfrac{1}{\sqrt{npq}}$, qu'en déduit-on sur la convergence de x_k ? Vérifier de même que $\dfrac{y_{k+1} - y_k}{x_{k+1} - x_k} = y_k \left(\dfrac{-x_k - \dfrac{q}{\sqrt{npq}}}{\dfrac{x_k q}{\sqrt{npq}} + 1 + \dfrac{q}{npq}} \right)$.

e. En admettant que lorsque n tend vers l'infini la suite $y_{k+1} - y_k$ tend vers dy, que $x_{k+1} - x_k$ tend vers dx, que x_k tend vers une valeur x et y_k vers une valeur y, on a alors : $\dfrac{dy}{dx} = -xy$, soit l'équation différentielle du début !

Conclure alors que y_k peut se représenter par une fonction de la forme $K.F$ du 2.f. (K est une constante que l'on ne cherche pas) et donner une expression de v_k.

Solution

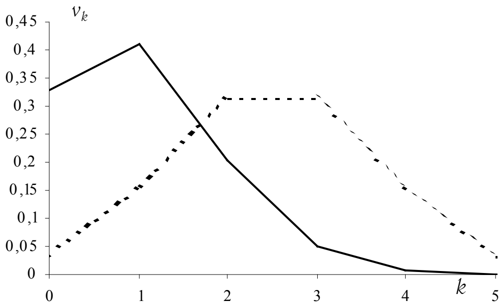
1. $F(x) = e^{-\left(\frac{x-a}{b}\right)^2}$; calculons les dérivées :

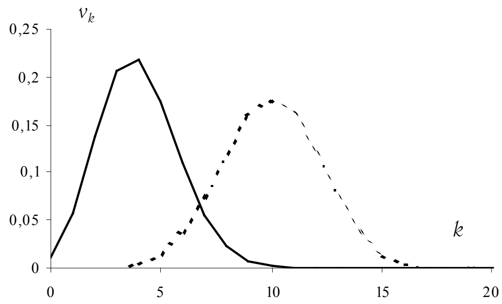
$$F'(x) = -2\frac{1}{b}\left(\frac{x-a}{b}\right)e^{-\left(\frac{x-a}{b}\right)^2},\ F''(x) = -2\frac{1}{b^2}\left(1 - 2\left(\frac{x-a}{b}\right)^2\right)e^{-\left(\frac{x-a}{b}\right)^2}\ ;$$

on a donc le maximum en $x_0 = a$ et $F(a) = 1$; quand aux points d'inflexion ils sont obtenus pour $1 = 2\left(\dfrac{x-a}{b}\right)^2 \Leftrightarrow \dfrac{x-a}{b} = \pm\dfrac{1}{\sqrt{2}} \Leftrightarrow x = a \pm \dfrac{1}{\sqrt{2}}b$.

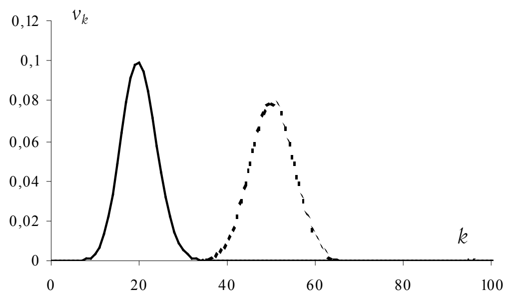
La distance entre les deux points est $|x_1 - x_2| = \left|\dfrac{2}{\sqrt{2}}b\right| = b\sqrt{2}$; pour que cette distance vaille 1, il faut $b = \dfrac{\sqrt{2}}{2}$.

2. a. Nous avons représenté la suite v_k pour $n = 5$, 20 et 100 ; en trait plein lorsque $p = 0{,}2$ et en pointillé lorsque $p = 0{,}5$.





On constate que plus n augmente plus la courbe ressemble à la fonction de Gauss mis à part le fait qu'elle soit décentrée. Comme la loi binomiale a pour moyenne np, cette quantité dans le cas où $n = 100$ vaut 20 pour $p = 0{,}2$ et 50 pour $p = 0{,}5$; on voit bien que cela correspond au maximum de chaque courbe.



Pour retrouver la fonction de Gauss il faut donc faire une translation de np vers la gauche.

L'autre paramètre important est l'écart-type, qui vaut pour la loi binomiale $\sigma = \sqrt{np(1-p)}$ qui correspond à l'écart entre les points d'inflexion ; il faudra alors faire en plus une homothétie sur les x pour ajuster correctement.

Reprenons $v_k = \begin{pmatrix} n \\ k \end{pmatrix}(1-p)^k p^{n-k}$: la somme des termes est le développement du binome de Newton avec $a = p$, $b = 1 - p$, on a donc $\displaystyle\sum_{k=0}^{n} v_k = (p + 1 - p)^n = 1^n = 1$.

b. On a $\dfrac{\begin{pmatrix} n \\ k+1 \end{pmatrix}}{\begin{pmatrix} n \\ k \end{pmatrix}} = \dfrac{\dfrac{n(n-1)(n-2)\dots(n-k+1)(n-k)}{(k+1)k(k-1)\dots 2.1}}{\dfrac{n(n-1)(n-2)\dots(n-k+1)}{k(k-1)\dots 2.1}} = \dfrac{n-k}{k+1}$ après simplification (il vaut mieux utiliser cette écriture sinon les calculs sont très pénibles).

Pour simplifier l'écriture notons $\sigma = \sqrt{npq}$: $y_k = \sigma v_k$, $y_{k+1} - y_k = (v_{k+1} - v_k)\sigma$; remarquons alors que $y_{k+1} - y_k = y_k\left(\dfrac{y_{k+1}}{y_k} - 1\right) = y_k\left(\dfrac{v_{k+1}}{v_k} - 1\right) = y_k\left(\dfrac{np - k - q}{(k+1)q}\right)$; par ailleurs

$$x_{k+1}-x_k=\frac{1}{\sqrt{npq}}=\frac{1}{\sigma} \text{ d'où } \frac{y_{k+1}-y_k}{x_{k+1}-x_k}=\sigma y_k\left(\frac{np-k-q}{(k+1)q}\right).$$

Simplifions maintenant l'intérieur de la parenthèse donnée dans l'énoncé :

$$\frac{-x_k-\dfrac{q}{\sigma}}{\dfrac{x_kq}{\sigma}+1+\dfrac{q}{\sigma^2}}=\frac{-\dfrac{k-np}{\sigma}-\dfrac{q}{\sigma}}{\dfrac{(k-np)q}{\sigma^2}+1+\dfrac{q}{\sigma^2}}=\frac{\dfrac{-k+np-q}{\sigma}}{\dfrac{kq-npq+\sigma^2+q}{\sigma^2}}=\sigma\frac{(np-k-q)}{(k+1)q}\ ;$$

la relation est prouvée.

c. $\frac{v_{k+1}-v_k}{v_k}=\frac{v_{k+1}}{v_k}-\frac{v_k}{v_k}=\frac{n-k}{k+1}\frac{p}{q}-1=\frac{np-kp-kq-q}{(k+1)q}=\frac{np-k-q}{(k+1)q}$. On retrouve au numérateur le terme $np-k-q$ dont le signe donne le sens de variation de v_k.

d. $x_k=\frac{k-np}{\sqrt{npq}}$; $x_{k+1}-x_k=\frac{k+1-np}{\sqrt{npq}}-\frac{k-np}{\sqrt{npq}}=\frac{1}{\sqrt{npq}}$; quand n tend vers l'infini ce terme tend vers 0 donc x_k converge (la distance entre deux termes consécutifs tend vers 0).

e. $x_k=\frac{k-np}{\sqrt{npq}}$ tend donc vers une limite x et $y_k=\sigma v_k$ tend vers une limite y, solution de $y'=-xy$, on a alors $\frac{y'}{y}=-x\Rightarrow \ln|y|=-\frac{1}{2}x^2+C\Rightarrow y=Ke^{-\frac{1}{2}x^2}=Ke^{-\left(\frac{x}{\sqrt{2}}\right)^2}$ qui est bien de la forme requise avec a = 0 et $b=\sqrt{2}$. Si on revient à v_k il faut remplacer x et y par x_k et y_k (étant sous-entendu qu'on est à la limite) :

$$\sigma v_k=Ke^{-\left(\frac{k-np}{\sigma\sqrt{2}}\right)^2}\Rightarrow v_k=K\frac{1}{\sigma}e^{-\left(\frac{k-np}{\sigma\sqrt{2}}\right)^2}.$$

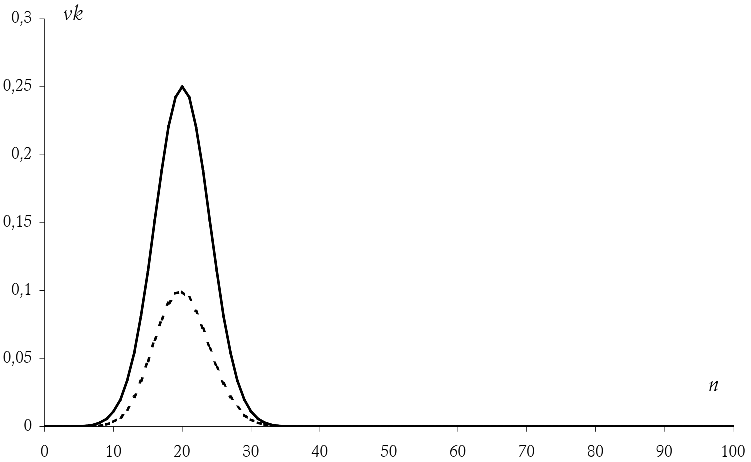
Pour calculer la valeur de K il faut utiliser $\sum_{k=0}^{n}v_k=1$ (ce qui est quasiment impossible) ou (en repassant en x et y) trouver K de sorte que : $K\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx=1$ (ce qui n'est pas simple non plus mais qui se fait).

$$\frac{v_{k+1}}{v_k}=\frac{\binom{n}{k+1}p^{k+1}q^{n-(k+1)}}{\binom{n}{k}p^kq^{n-k}}=\frac{n-k}{k+1}\frac{p^{k+1}}{p^k}\frac{q^{n-k-1}}{q^{n-k}}=\frac{n-k}{k+1}\frac{p}{q}.$$

On a $v_{k+1}>v_k$ si $\frac{n-k}{k+1}\frac{p}{q}>1\Leftrightarrow pn-pk>qk+q\Leftrightarrow k(q+p)<pn-q\Leftrightarrow k<pn+p-1$.

$p-1$ est forcément petit : au plus il vaut 0, au moins il vaut -1 ; on a donc $pn-1<k<pn$.

Si on prend l'entier le plus proche de pn, c'est la valeur de k pour laquelle v_k change de sens de variation. En quelque sorte c'est la valeur de k pour laquelle la loi binômiale est maximale et v_{k_0} est son maximum (remarquez que np est la moyenne de la loi binômiale).



Ci-dessus sont tracés les v_k d'origine (loi binomiale, n = 100, p = 0,2) en pointillés et les nouveaux v_k (loi de Laplace-Gauss) en trait plein.

La différence de hauteur provient bien sûr du facteur K que l'on obtient facilement en faisant la somme des v_k avec le tableur.

On trouve alors $K = 2,5066...$; élevons cette valeur au carré (à tout hasard bien sûr) : $K^2 = 6,2831...$

Depuis le temps que vous faites des maths vous avez reconnu tout de suite… 2π ! Eh oui, notre coefficient K vaut $\sqrt{2\pi}$.

Voilà notre petit voyage au pays de la loi de Gauss s'achève ici : cette loi représente alors une densité de probabilité pour une v.a. x, limite d'une loi binomiale B(n, p) ; sa moyenne est alors $\overline{x} = np$, son écart-type est $\sigma = \sqrt{npq}$ et on calcule la probabilité de l'événement [a ; b] en faisant

$$P(a \le x \le b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\overline{x})^2}{2\sigma^2}} dx .$$

On ne sait pas intégrer la fonction de Gauss, aussi les valeurs sont données par des tables ou par des fonctions dans le tableur (LOI.NORMALE(…))

Cette loi se retrouve fréquemment car il existe un théorème (*théorème de la limite centrale*) disant que si vous prenez une suite de v.a. x_n de loi <u>quelconque</u> indépendantes, de moyenne $\overline{x}$ et d'écart-type σ alors la moyenne arithmétique de ces v.a. : $\frac{x_1 + x_2 + ... + x_n}{n}$ tend vers une loi normale de moyenne $\overline{x}$ et d'écart-type $\frac{\sigma}{\sqrt{n}}$.

Par exemple si chaque v.a. représente la taille d'un individu, la répartition de ces tailles suivra une loi normale.