Atelier autour des probabilités et de la statistique
Journée de formation aux nouveaux programmes de terminales

Octobre 2012

Loi normale et calcul de probabilité

Dans une fabrique de peintures décoratives, une machine remplit automatiquement avec de la peinture des bidons de 51 centilitres.

Pour pouvoir être commercialisé, un bidon doit contenir au moins 48 centilitres de peinture.

La quantité de peinture en centilitres fournie par la machine peut être modélisée par une variable aléatoire X qui suit une loi normale de moyenne μ et d'écart-type $\sigma = 1,2$.

1) La machine est réglée sur $\mu = 50$.

   a) Quel est le pourcentage de bidons qui ne pourront pas être commercialisés ?

   b) Quel est le pourcentage de bidons qui risquent de déborder au moment du remplissage ?

   c) Quelle devrait être la contenance, au millilitre près, des bidons pour que moins de 10 % des bidons ne débordent ?

2) Réglage de la machine.

Quelle devrait être la quantité moyenne maximale de remplissage de sorte que moins de 10 % des bidons ne débordent ?

Intervalle de fluctuation asymptotique et prise de décision

Le riz trois couleurs est vendu dans le commerce et contient plusieurs types de grains, dont les proportions sont connues. Ainsi, les grains de couleur marron sont annoncés comme représentant 20% de l'ensemble des grains.

Les élèves d'une classe de Terminale ont voulu vérifier cette information. Pour cela, ils ont choisi d'observer un échantillon aléatoire de grains de riz et de construire un intervalle de fluctuation asymptotique au seuil de 95% pour la proportion de riz marron.

Ils ont donc constitué un échantillon, que l'on peut considérer aléatoire, de 690 grains de riz. Ils ont dénombré 140 grains de riz marron.

La règle de décision est la suivante : si la proportion de grains de riz marron dans l'échantillon n'appartient pas à l'intervalle de fluctuation, on rejette l'hypothèse selon laquelle les grains de riz marron représentaient 20% des grains de riz pendant la période où les paquets de riz trois couleurs utilisés pour l'expérience ont été vendus.

1) Déterminer l'intervalle de fluctuation asymptotique J au seuil de 95% pour la proportion de grains de riz marron.

2) Calculer la proportion de grains de riz marron dans l'échantillon. Que peut-on en conclure ?

3) Dans le même échantillon, il y avait 152 grains de riz noir pour une proportion annoncée de 20% et 125 grains de riz rouge, pour une proportion annoncée de 10%. Que peut-on conclure de ces résultats ?

4) Programmer un algorithme qui permet d'automatiser les calculs précédents en s'appuyant sur l'intervalle de fluctuation asymptotique au seuil de 95 %. Cet algorithme devra permettre d'accepter ou de rejeter l'hypothèse d'une probabilité p supposée (ici par le négociant en riz) sur la réalisation d'un événement suite à l'obtention d'une fréquence f observée sur un échantillon de taille n.

De la loi binomiale à la loi normale par le tableur

Partie A

Une variable aléatoire X_n suit une loi binomiale de paramètres n et p.
Créer une feuille de calcul sur un tableur de type OpenOffice comme ci-dessous permettant de déterminer la loi de probabilité de X_n suivant n et p.

	A	B	C	D
1	n=	50	p=	0,4
2				
3	***k***	***P(Xn = k)***		
4	0	8,0828E-012		**Espérance**
5	1	2,6943E-010		20
6	2	4,4006E-009		
7	3	4,6940E-008		
8	4	3,6770E-007		**Ecart-type**
9	5	2,2552E-006		3,464101615
10	6	1,1276E-005		

1) Compléter la feuille de calcul de l'espérance m de X_n et de son écart-type s.

2) Représenter ces deux premières colonnes par un histogramme, en réalité un diagramme en bâtons pour lequel vous aurez réduit l'espacement entre les bâtons de façon à simuler un histogramme.

Partie B

On considère la variable aléatoire $Y_n = \dfrac{X_n - m}{s}$.

On souhaite maintenant construire l'histogramme "normalisé" de Y_n, c'est-à-dire pour lequel la somme des aires des rectangles sera égale à 1.

1) Quel est l'écart entre deux valeurs consécutives de Y_n ?

2) Augmenter la feuille de calcul de deux nouvelles colonnes :
   - celle des valeurs y_k pour k entre 0 et n prises par Y_n ;
   - celle des valeurs des hauteurs des rectangles ayant pour aire $P(Y_n = y_k)$ et pour largeur l'écart entre deux valeurs consécutives de Y_n.

3) Représenter ces deux dernières colonnes par un histogramme. Faire varier n et p. Que remarquez-vous ?

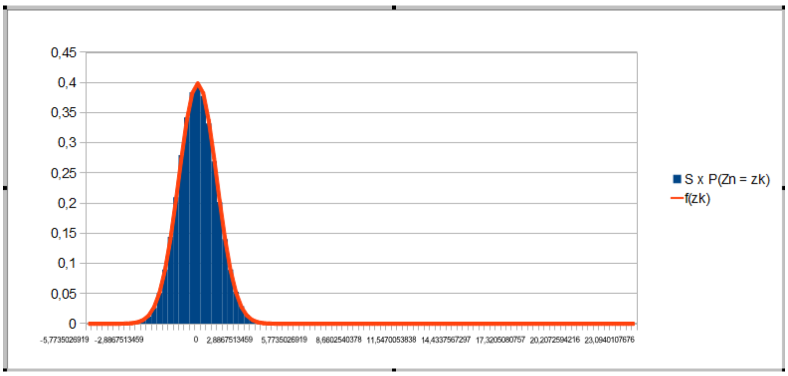
Partie C

On introduit la fonction f définie sur IR par $f(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$.

1) Créer une nouvelle colonne à la suite de la précédente contenant les images des valeurs de y_k par cette fonction.

2) Il s'agit d'insérer la courbe de cette fonction sur l'histogramme représentant Y_n :

Modifier la plage de données du diagramme précédent afin d'insérer les valeurs des images $f(y_k)$. Vous devriez obtenir lors de cette étape deux histogrammes.
Modifier ensuite le type de diagramme en choisissant le type "Colonne et ligne" pour rendre le diagramme mixte.
Que remarquez-vous ? Cette remarque subiste-t-elle si vous faites varier n et p ?



On admet qu'en diminuant progressivement les amplitudes des intervalles de l'histogramme, les rectangles constituant cet histogramme seront de bases presque nulles et toutes égales. Les "sommets" de ces rectangles seront alors des points tellement rapprochés qu'ils formeront une courbe, dite **courbe de tendance de l'histogramme**".
On règle les valeurs n sur 50 et p sur 0,4.

3) Expliquer pourquoi $P(-1 \le Y_n \le 2) = P(17 \le X_n \le 26)$.

4) Calculer, à l'aide de votre calculatrice, l'aire sous la courbe de la fonction f entre $x = -1$ et $x = 2$.
   Interpréter graphiquement cette intégrale et en déduire une valeur approchée de $P(17 \le X_n \le 26)$.

Prolongement possible : comment réaliser cette même activité à l'aide de Geogebra et de son tableur ?

Loi uniforme en géométrie

Le plan est muni d'un repère orthonormé direct (O , I , J) .
Γ est le quart de cercle de centre O de rayon 1 d'origine I d'extrémité J.

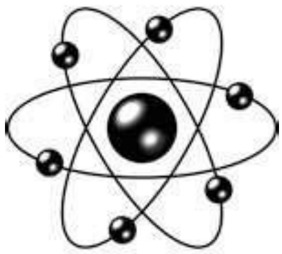
M est un point choisi au hasard sur Γ.

Quelle est la probabilité pour que le triangle JOM ait un périmètre supérieur ou égal à 3 ?

Mise en œuvre : vous procéderez par simulation et contrôlerez votre résultat par des considérations géométriques dans l'un des trois cas suivants :

- **Cas 1 :** M est le point associé au réel t choisi au hasard dans $[0 ; \pi / 2]$.
- **Cas 2 :** M a pour abscisse x choisie au hasard dans [0 ; 1].
- **Cas 3 :** M a pour ordonnée y choisie au hasard dans [0 ; 1].

Désintégration radioactive et simulation



Un nombre N_0 d'atomes radioactifs identiques ont chacun la probabilité p de se désintégrer au cours d'une unité de temps.

Objectif : simuler ces désintégrations, et tracer l'histogramme des fréquences de désintégration par unité de temps.

Algorithme de simulation

Dans le morceau d'algorithme ci-contre :

- la variable `N` représente le nombre d'atomes radioactifs à un instant donné ;
- `Nnew` compte les atomes « survivants » après une unité de temps ;
- p est la probabilité de désintégration d'un atome.

```
POUR compteur ALLANT_DE 1 A N
   DEBUT_POUR
   SI (random()<1-p) ALORS
      DEBUT_SI
      Nnew PREND_LA_VALEUR Nnew+1
      FIN_SI
   FIN_POUR
```

Compléter cet algorithme de manière à tracer l'histogramme des fréquences de désintégration par unité de temps, sur une durée à votre convenance.