Lois conjointes
Théorèmes limites
ESSI1

Exercice 5.1

Dans un certain village

- *15% des familles n'ont pas d'enfants*
- *20% des familles ont un enfant*
- *35% des familles ont deux enfants*
- *30% des familles ont trois enfants*
- *Chaque enfant a une chance sur deux d'être une fille indépendamment du sexe de ses éventuels frères et sœurs*

Soit F et G les deux variables aléatoires représentant le nombre de filles et de garçons dans une famille de ce village.
Déterminer la loi conjointe de F et G.

j / i	0	1	2	3	P(G=i)
0	**15/100**	**10/100**	**35/400**	**30/800**	**30/80**
1	**10/100**	**35/200**	**90/800**	**0**	**31/80**
2	**35/400**	**90/800**	**0**	**0**	**16/80**
3	**30/800**	**0**	**0**	**0**	**3/80**
P(F=j)	**30/80**	**31/80**	**16/80**	**3/80**	**1**

Exercice 5.2
On réalise n+m épreuves indépendantes, chacune ayant probabilité de succès n.
X= nombre de succès pendant les m premières épreuves
Y= nombre de succès pendant les n dernières épreuves
Z= nombre total de succès

X et Y sont elles indépendantes ?
X et Z sont elles indépendantes ?

P(X=i,Y=j)=$C_m^i p^i (1-p)^{m-i} C_n^i p^j (1-p)^{n-j}$ =P(X=i)+P(Y=j) donc X et Y sont indépendantes.

P(X=i,Z=j)=0 si j<i Or P(Z=j) et P(X=i) sont non nulles

Exercice 5.4
Montrez que si X et Y sont deux variables aléatoires discrètes
Cov(X,Y)=Cov(Y,X)
Cov(X,X)=Var(X)
Cov(aX,Y)=aCov(X,Y)
Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)

Exercice 5.5 :
Montrez que si les Xi sont deux à deux indépendantes, alors alors la variance de leur somme est égale à la somme de leur variance.

$$Var(\sum_{i=1}^{n} X_i) = Cov(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i)$$

$$= \sum_{i=1}^{n} \sum_{i=1}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq j} \sum Cov(X_i, X_j)$$

Exercice 5.6
Prouver l'inégalité de Tchebychev

Il suffit d'appliquer l'inégalité de Markov avec a=k2 à la variable aléatoire (X-m)2.
$P((X-m)^2 \geq k^2) \leq E[(X-m)^2]/k^2 = \sigma^2/k^2$

Or $P((X-m)^2 \geq k^2) = P(|X-m| \geq k)$

Exercice 5.7

Supposons que je produise en moyenne 50 diapositives par jour, quelle est, au plus, la probabilité pour que demain je produise plus de 75 diapositives ?
Si je sais de plus que la variance de ma production journalière est de 25, quelle est, au moins, la probabilité pour que demain je produise entre 40 et 60 diapositives ?

D'après l'inégalité de Markov
$P(X \geq 75) \leq E[X]/75=2/3$
D'après l'inégalité de Tchebychev
$P(|X-50| \geq 10) \leq 1/4$
Donc $P(|X-50| < 10) > 3/4$
Remarque:
Markov: $P(X \geq 60) \leq 5/6$
Tchebychev: $P(X \geq 60) = (P(|X-50| \geq 10))/2 \leq 1/8$

Exercice 5.8

Soit X une variable aléatoire de moyenne m et d'écart-type s. Le rapport r=|m|/s est appelé rapport signal-bruit. On définit D=|(X-m)/m|, l'erreur relative de X par rapport à son signal m.
Montrez que pour tout a >0 P(D < a) ≥1- 1/(ra)2

Il suffit d'appliquer l'inégalité de Tchebychev :
$P(|X-m| \geq k) \leq \sigma^2/k^2$
$P(D < a) = P(|(X-m)/m| < a) =$
$P(|X-m| < |m|a) = 1 - P(|X-m| \geq |m|a)$

Or $P(|X-m| \geq |m|a) \leq \sigma^2/(|m|a)^2 = 1/(ra)^2$

Donc $P(D < a) \geq 1- 1/(ra)^2$

Exercice 5.9

Supposons que X suive une loi normale d'espérance m et de variance σ^2, calculez la borne donnée par l'inégalité de Tchebychev sur P(|X-m| ≥ 2σs) ainsi que la valeur exacte de cette probabilité, comparez.

L'inégalité de Tchebychev donne
$P(|X-m| \geq 2\sigma) \leq \sigma^2/4\sigma^2 = 1/4$

Alors que la loi exacte donne
$P(|X-m| \geq 2\sigma) = P(|X-m|/\sigma \geq 2) = P((X-m)/\sigma \geq 2) + P((X-m)/\sigma \leq -2) =2(1-F(2)) \approx 0{,}0456$

Exercice 5.10

On effectue n parties de pile ou face. Trouver n pour que l'on puisse affirmer que la fréquence d'apparition de pile durant ces n lancers soit comprise entre 0.45 et 0.55 avec une probabilité supérieure à 0.90.

La variable aléatoire X qui compte le nombre de piles parmi les n épreuves est la variable binomiale de paramètre (n,1/2).
La fréquence F d'apparition de pile est X/n et :E[F] = E[B(n,1/2)]/n = n/2n = ½
V[F] = V[B(n,1/2)/n] = n/4n2 = 1/4n
L'inégalité de Tchebychev donne
P(|F-1/2| ≥ k) ≤ V[F]/k2 = 1/4nk2

On cherche n pour que P(|F-1/2| ≤ k) ≥ 0. Il suffit donc que 1/4nk2 ≤ 0.10 avec k = 0.05 soit n > 1000.

Exercice 5.11:

Un astronome désire mesurer la distance, en années-lumière entre la terre et une certaine étoile. Chaque fois qu'il réalise une mesure, il obtient une distance approchée du fait de petites erreurs dues entre autres, aux influences atmosphériques. Il prévoit donc de faire plusieurs mesures et de prendre la moyenne de ces mesures comme estimation de la distance réelle. Il suppose que les valeurs mesurées sont des variables aléatoires indépendantes suivant toutes la même loi, d'espérance commune d (la vraie valeur) et de variance commune 4. Combien de mesures doit-il effectuer pour être raisonnablement sûr (95%) que l'erreur soit inférieure à une demi-année-lumière ?

Notons Xi la i-ème mesure, on cherche : $P(-0,5 \le \frac{\sum_{i=1}^{n} X_i}{n} - d \le 0,5)$

Le théorème central limite établit que $Z_n = \frac{\sum_{i=1}^{n} X_i - nd}{2\sqrt{n}}$

suit approximativement une loi normale standard.
Donc en se ramenant à la loi normale standard :

-

$$P(-0,5 \le \frac{\sum_{i=1}^{n} X_i}{n} - d \le 0,5) = P(-0,5 \le \frac{\sum_{i=1}^{n} X_i - nd}{n} \le 0,5)$$

$$P(-0,5\frac{\sqrt{n}}{2} \le Z_n \le 0,5\frac{\sqrt{n}}{2})$$

$$\approx 2\Phi(\frac{\sqrt{n}}{4}) - 1$$

S’il veut que l’erreur soit d’au plus une demi-année lumière avec probabilité 0,95 il devra choisir n tel que

$$2\Phi(\frac{\sqrt{n}}{4})-1\geq 0{,}95$$

Donc n > 61.

Exercice 5.13

On fait une suite de n tirages Pile ou Face avec une pièce non biaisée (p=1/2). En utilisant le théorème de la limite centrale, calculer une approximation de la probabilité

- *que le nombre de Pile soit dans l’intervalle [45..55] si on fait 100 tirages*
- *que le nombre de Pile soit dans l’intervalle [450..550] si on fait 1000 tirages*

Combien faut-il faire de tirages pour que le nombre de Pile se trouve dans l’intervalle [0.45n ,0.55n] avec une probabilité de 95% ?
Comparer avec la borne obtenue par Tchebychev.

Il s’agit de lois binomiales que l’on va approximer par des lois normales.
Considérons la suite de variables aléatoires
Xi = 1 si la pièce tombe sur pile au ième tirage et Xi = 0 sinon.
Soit Sn= (X1+…+Xn)
E[Xi] = 1/2 et Var(Xi) = pq = ¼
D’après le théorème de la limite centrale $\frac{X_1+....+X_n-nm}{\sigma\sqrt{n}}$ est approximée par N

La fonction de répartition de (Sn - nm)/sn1/2 tend vers une loi normale standard quand n tend vers l’infini.
Avec m = 1/2 et s = 1/2
Soit (Sn - n/2)/(n1/2/2)

(Sn - n/2)/(n1/2/2) approximée par N

Pour n=100, on a :
P(45 ≤ S100 ≤ 55) = P((45-50)/5 ≤ (S100 -50)/5≤ (55-50)/5)
P(-1 ≤ N ≤ 1) =2F(1)-1 = 0,68

Pour n=1000, on a :
P(450 ≤ S1000 ≤ 550) ≈ 2F(√10)-1 = 0,999
(car n/2 = 500 et (550-500)2/√1000 = √10)
(Sn - n/2)/(n1/2/2) approximée par N
P(0,45n≤ Sn ≤0,55n) =
P(0,45n - 0.50n≤ Sn - 0.50n ≤0,55n - 0.50n) =
P(-0,05n/(√n 1/2) ≤ (Sn - 0.50n)/(√n 1/2) ≤0,05n/ (√n 1/2)
2F(0.05n/(√n 1/2))-1 = 2F(0,1√n)-1

Pour que cette valeur soit supérieure à 0,95 , on doit avoir n > 400 environ.
A comparer avec la borne obtenue par Tchebychev : n > 1000 assure une fiabilité supérieure à 0,90.