TD 2 : Fonction de répartition et théorèmes asymptotiques

Exercice 1 (Simulation d'une loi à partir de la loi uniforme)**.** *Soit Y une variable aléatoire réelle à valeur dans l'intervalle I. On suppose que la fonction de répartition F_Y de la variable Y est continue sur $\mathbb{R}$ et est strictement croissante sur I de telle sorte que la bijection réciproque $F_Y^{-1} :]0,1[\to I$ soit bien définie.*

1. *Soit U une loi uniforme $\mathcal{U}(0,1)$. Montrer que la variable aléatoire réelle $X = F_Y^{-1}(U)$ a la même loi que la variable Y (i.e montrer que la fonction de répartition de la variable aléatoire X est F_Y).*
2. *Soit $\lambda > 0$. À l'aide de la question précédente, montrer que la variable aléatoire $X = -\frac{1}{\lambda}\ln(1-U)$ suit une loi exponentielle $\mathcal{E}(\lambda)$.*
3. *En utilisant le logiciel RStudio, on simule un échantillon expérimental $x = (x_1, \ldots, x_{100})$ de taille 100 d'une loi exponentielle $\mathcal{E}(2)$ via le code suivant :*

```
u=runif(100)
x=-log(1-u)/2
```

4. *On compare la fonction de répartition aléatoire de l'échantillon expérimental x définie par*

$$\forall t \in \mathbb{R} \quad F_{100}(t) = \frac{1}{100}\sum_{i=1}^{100} \mathbf{1}(x_i \leq t),$$

à la fonction de répartition de la loi $\mathcal{E}(2)$ via le code suivant :

```
xord=sort(x)
y=(1:100)/100
plot(xord,y,type="s")
curve(pexp(x,2),add=TRUE,col="red")
```

5. *Modifier le code pour un échantillon expérimental de taille* 1000 *d'une loi exponentielle $\mathcal{E}(0.5)$.*

Exercice 2. *On lance mille fois une pièce parfaitement équilibrée. Quel est la probabilité d'avoir au moins 545 piles. Calculer exactement cette probabilité puis donner une approximation de cette probabilité.*

Exercice 3. *Dans son dernier souffle, César a prononcé la phrase "Tu quoque mi fili". On estime qu'en prononçant cette phrase César a expiré un litre d'air. Donner un ordre de grandeur de la probabilité qu'une personne inspirant un litre d'air avale au moins une molécule ayant été expiré par César lors de son dernier souffle. Données : un litre d'air contient $2,7 \times 10^{22}$ molécules et l'atmosphère terrestre contient environ $1,1 \times 10^{44}$ molécules.*

Exercice 4. *On note p la proportion de personnes positives à un virus dans une population (très grande). Pour n personnes choisies au hasard dans la population, on pose S le nombre de personnes positives au virus et on modélise la loi de S par une loi Binomiale $\mathcal{B}(n,p)$. Mille personnes se réunissent dans une salle pour assister à un concert (donc on prend $n = 1000$). Lorsque $p \in \{0.002, 0.01\}$, calculer la probabilité $\mathbb{P}(S \geq 5)$ de façon exacte, en approchant la loi de S par une loi de Poisson et en approchant la loi de S par une loi normale. Que remarquez-vous ?*

Exercice 5. *On suppose que X suit une loi normale de loi $\mathcal{N}(25, 25)$.*

1. *Calculer les probabilités :* $\mathbb{P}(X \geq 30)$, $\mathbb{P}(X \leq 15)$, $\mathbb{P}(18 \leq X \leq 32)$.
2. *Déterminer la valeur de a telle que :* $\mathbb{P}(X \leq a) = 0.95$.
3. *Donner les bornes d'un intervalle de fluctuation pour X centré en 25 de niveau 99%.*
4. *On suppose maintenant que X suit une loi $\mathcal{N}(\mu, \sigma^2)$ avec $\mu \in \mathbb{R}$ et $\sigma > 0$ quelconques. Calculer :* $\mathbb{P}(\mu - k\sigma \leq X \leq \mu + k\sigma)$ *pour $k = 1, 2, 3$ et 4.*

Exercice 6. *Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes et de même loi d'espérance $\mu \in \mathbb{R}$ et de variance σ^2. On note $Y = X_1 + \cdots + X_n$.*

1. *Exprimer $\mathbb{E}(Y)$ et $\mathrm{var}(Y)$ en fonction de n, μ et σ^2.*
2. *La charge maximale du téléphérique de l'aiguille du midi est de 5040 kg; On admet que le poids d'un passager choisi au hasard se modélise par une loi normale de moyenne 70 kg et d'écart type $8kg$. On pose Y le poids total aléatoire de n personnes montant ensemble à bord du téléphérique.*
   (a) *Préciser la loi de probabilité de la variable aléatoire réelle Y.*
   (b) *Calculer la probabilité d'avoir une surcharge lorsque 75 personnes montent à bord .*
   (c) *Quel est le nombre maximum de personnes à autoriser à bord pour que le risque de surcharge soit inférieur à 1% ?*

Exercice 7. *Soit $U_1, \ldots, U_n$ une suite de variable aléatoire indépendantes de loi uniforme $\mathcal{U}(0, 1)$. La loi de Irwin-Hall est la loi de la variable aléatoire $X = U_1 + \ldots + U_n$.*

1. *Montrer la convergence en loi*

$$\forall t \in \mathbb{R} \quad \lim_{n \to +\infty} \mathbb{P}\left(\frac{X - n/2}{\sqrt{n/12}} \leq t\right) = \underbrace{\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx}_{=\Phi(t)}.$$

2. *On pose $Z = (X - n/2)/\sqrt{n/12}$. Le code R suivant permet de simuler un échantillon expérimental $z = (z_1, \ldots, z_{100})$ de taille 100 de la variable Z lorsque $n = 12$.*

```
z=numeric(100)
for(i in (1:100))
{
  z[i]=sum(runif(12))-6
}
```

3. *On compare la fonction de répartition aléatoire de l'échantillon expérimental z définie par*

$$\forall t \in \mathbb{R} \quad F_{100}(t) = \frac{1}{100} \sum_{i=1}^{100} \mathbf{1}(z_i \leq t),$$

   à la fonction de répartition de la loi $\mathcal{N}(0, 1)$ via le code suivant :

```
zord=sort(z)
y=(1:100)/100
plot(zord,y,type="s")
curve(pnorm(x),add=TRUE,col="red")
```

4. *Modifier le code en prenant $n \in \{3, 12\}$ et en augmentant la taille de l'échantillon expérimental de Z à 1000.*

Contrairement à la fonction de répartition de la loi normale centrée réduite Φ, la fonction de répartition F_X de la loi de Irwin-Hall est explicite :

$$\forall t \in [0,n] \quad F_X(t) = \sum_{k=0}^{\lfloor t \rfloor} \frac{(-1)^k}{k!(n-k)!}(t-k)^n.$$

La convergence en loi établie à la question 1 donne l'approximation : $\Phi(t) \approx F_X\left(n/2 + t\sqrt{n/12}\right)$.

5. *Le code R suivant permet de comparer graphiquement ϕ et son approximation :*

```
F_X=function(t,n)
{
  if((t>0)&&(t<n))
  {
    s=0
    for(k in (0:floor(t)))
    {
      s=s+(-1)^k/(factorial(k)*factorial(n-k))*(t-k)^n
    }
  }
  else if(t<0){s=0} else{s=1}
  return(s)
}

t=seq(-6,to=6,by=0.01)
taille=length(t)
y=numeric(taille)
n= #entrer une valeur pour n
for(i in (1:taille))
{
  y[i]=F_X(sqrt(n/12)*t[i]+n/2,n)
}

plot(t,y,type="l")
curve(pnorm(x),add=TRUE,col="red")
```

Comparer Φ et son approximation pour $n \in \{1, 3, 12\}$.

Exercice 8. *Soit X une variable aléatoire à valeur dans $\mathbb{Z}$ telle que pour tout $i \in \mathbb{Z}$ on a $\mathbb{P}(X = i) > 0$ et Y une variable aléatoire à valeur dans $\mathbb{N}^*$ indépendante de X telle que pour tout $i \in \mathbb{N}^*$ on a $\mathbb{P}(Y = i) > 0$. On pose $W = X/Y$. Quelle spécificité, un peu pathologique, a la fonction de répartition de la variable aléatoire réelle W ?*