CPP-la prépa des INP 2 ème année
année 2015-2016

TP 2: Probabilités et statistiques

Ne pas oublier de charger le chemin pour *stixbox*. Aller dans l'onglet *file* puis *set path* et ajouter le chemin

```
C:\Program Files\MATLAB\R2012a\toolbox\shared\stixbox
```

Partie 1. Intervalles de confiance

1) On va ici commencer par simuler des variables aléatoires indépendantes de loi de Bernoulli de paramètre p. Que fait la commande:

```
>> A= rand(1,5)
>> A<1/2
```

Rappeler la construction de l'intervalle de confiance asymptotique de niveau 0,95 pour le paramètre p dans un modèle d'échantillon $(X_1, \dots, X_n)$ de Bernoulli. Cette construction est basée sur le théorème central limite et fait appel à l'estimateur de p:

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Compléter le programme suivant, où l'on tracera le comportement de $\hat{p}_n$ ainsi que celui de l'intervalle de confiance associé à p lorsque la taille de l'échantillon grandit. On tracera également la valeur théorique de p sur le graphique.

On pourra également tracer l'intervalle de confiance asymptotique au niveau 0,99.

```
function A=IntConf(n,p)
U=rand(1,n);
B=...;  % loi de Bernoulli Ber(p)
l=...;      %taille de l'intervalle de confiance
C=[1:n];
A=...;  % A(n) estimateur p_n de p
plot(C,A,'b',C,p,'y', C,..., 'r')  % intervalle de confiance à tracer!
```

2) On va maintenant vérifier que pour n suffisamment grand, la probabilité de tomber dans l'intervalle de confiance de niveau 0,95 est effectivement (proche) de 0,95. Pour cela, on se fixe n suffisamment grand (pour

pouvoir utiliser le TCL), et on construit un grand nombre N d'intervalles de confiance pour p. On calculera alors la proportion de fois où le vrai paramètre p appartient à l'intervalle de confiance obtenu.

```
function prop=IC(n,N,p)
NbIC=0;
l=...;      %taille de l'intervalle de confiance
for i=1:N
B=...;   % un échanttillon de loi Ber(p)
pn=...;  % estimateur de p
a=pn -l;
b=pn+ l;
if a<=p if p<=b
   NbIC=NbIC +1;
    end
end
end
prop=...
```

Partie 2. Tests statistiques

1) On se place encore dans le modèle d'échantillon de Bernoulli. On suppose que la taille de l'échantillon est de 100.

On souhaite tester l'hypothèse nulle:

$$H_0 : p = p_0$$

contre l' hypothèse alternative:

$$H_1 : p \neq p_0$$

et contre l'hypothèse alternative

$$H_1' : p > p_0$$

Pour le moment on prend $p_0 = 1/2$.

En cours, on a construit un test statistique de niveau 0,05 basé sur le comportement de la variable aléatoire

$$Y_n = 2\sqrt{n}\left(\hat{p}_n - \frac{1}{2}\right)$$

Rappeler la décision que donne le test en fonction de la valeur prise par Y_n. A l'aide d'une simulation , vérifier que pour les 2 tests, l'erreur de première espèce est bien de l'ordre 0,05, c'est à dire la probabilité de déclarer que H_0 est fausse alors que H_0 est vraie.

A l'aide d'une simulation, calculer la puissance du test pour différentes valeurs de p, c'est-à-dire la probabilité de rejeter H_0 lorsqu'effectivement $p \neq 1/2$. On prendra notamment $p = 0.2$, $p = 0.7$, $p = 0,52$.

2) Refaire la question précédente lorsque $p_0 = 0,3$

3) **Voir feuille de TD 3 exercice 1.**

On fait l'hypothèse que l'on observe $X_1, \ldots, X_n$ des variables aléatoires indépendantes de même loi $\mathcal{N}(\mu, \sigma^2)$ avec $\sigma = \frac{1}{8}$.

On veut tester l'hypothèse

$$H_0 : \mu = \mu_0$$

contre l'hypothèse alternative:

$$H_1 : \mu \neq \mu_0.$$

On considère ici que $\mu_0 = 1,76$. Ecrire le test précédent. Vérifier le niveau du test. Calculer (par simulation) la puissance du test pour des données de loi $\mathcal{N}(1,78, \frac{1}{8}^2)$ pour $n = 30, 100, 150$.

Calculer également la puissance du test pour des données de loi $\mathcal{N}(1,76, \frac{1}{4}^2)$, ainsi que des loi $\mathcal{E}(1/1,76)$.

On pourra aussi tester H_0 contre l'hypothèse alternative

$$H_1' : \mu > \mu_0$$

4) Refaire la question précédente en considérant que la variance σ^2 est inconnue: on utilisera alors le test de Student (**cf feuille td 3 exercice 2**).

Partie 3. **Estimateurs:** Comparer numériquement le risque quadratique des 2 estimateurs de $e^{-\theta}$ obtenus à l'exerice 4 de la feuille 1 dans le cas d'un échantillon de loi de Poisson de paramètre θ.

(On pourra utiliser la commande *poissrnd*.)