

Statistiques
Master Statistique et économétrie
Exercices Mise à Niveau - Feuille n^o 2

V. Monbet

Master 1 - 2012

L'objectif de ce TD est de voir comment on peut utiliser le logiciel R pour répondre à certaines questions de statistique et/ou vérifier des résultats théoriques.

1. Générer des échantillon de taille $n = 30$ selon les lois
 - de Poisson (`rpois`).
 - de Gauss centrée réduite (`rnorm`).
 - de Gauss de moyenne 5 et de variance 2.
 - uniforme entre -1 et 1 (`runif`).
 - de Bernouilli (utiliser la fonction (`runif`)).
2. Générer un échantillon selon une loi de Gauss de moyenne 5 et de variance 2 et représenter son histogramme (`hist`).
3. Estimer les moyennes et variances pour les échantillons précédents (`mean`, `var`, `sd`). Faire plusieurs tirages et comparer les statistiques estimées. Interpréter les résultats obtenus.
4. Loi des grands nombres.
Les commandes ci-dessous permettent d'illustrer la loi des grands nombres. Exécutez les puis interprétez, avec précision, le graphique obtenu.

```
Nexp = 100
nvec = seq(10,100,10)
nvec.len = length(nvec)
mx = matrix(0,Nexp,nvec.len)
k = 0
for (n in nvec) {
  k = k+1
  x = matrix(runif(n*Nexp),n,Nexp)
  mx[,k] = apply(x,2,mean)
}
boxplot(mx)
lines(0:(nvec.len+1),.5*matrix(1,1,nvec.len+2),lwd =2,col='red')
```

Adaptez ces lignes de code pour d'autres distributions puis d'autres tailles d'échantillon. Concluez.

5. TCL

Proposez une démarche similaire à celle de la question précédente afin d'illustrer le TCL.

6. Test de Student

Des chercheurs ont réalisé une expérience visant à mettre en évidence l'effet d'une séance d'intervention motivante brève sur le comportement relatif à la consommation d'alcool.

Soixante sujets, qui ont déclaré avoir bu occasionnellement à 2 reprises ou plus au cours du mois précédant l'expérience ont été affectés au hasard soit dans un groupe contrôle, sans traitement (31 sujets) soit dans un groupe expérimental dit "groupe d'intervention brève" (29 sujets).

Le comportement des sujets est mesuré par la variable "nombre de verres bus par semaine". Les sujets sont évalués avant l'expérience (condition de référence). Chacun des sujets du groupe d'intervention brève bénéficie d'un entretien personnalisé relatif aux problèmes liés à l'alcool. Six semaines après l'entretien, l'ensemble des sujets est de nouveau évalué.

Les données sont stockées dans le fichier `Intv-breve.csv` disponible sur la page du cours (<http://>).

Comparer les moyennes dans les groupes de référence :

- (a) Charger les données

```
IB = read.csv2("Interv-breve.csv")
```

et visualiser une version synthétique du contenu de la table IB.

```
str(IB)
```

- (b) Calculer les statistiques descriptives.

```
summary(IB)
```

- (c) Comparer les moyennes dans les groupes de référence (en supposant l'égalité des variances)

```
t.test(IB[,1], IB[,2], var.equal=TRUE)
```

Interpréter les résultats.

- (d) Comparer les moyennes avant et après traitement

```
t.test(IB[,1], IB[,3], var.equal=TRUE, paired=TRUE)
```

```
t.test(IB[,2], IB[,4], var.equal=TRUE, paired=TRUE)
```

Interpréter les résultats.

- (e) La fonction `t.test` permet aussi d'obtenir des intervalles de confiance (à 95% par défaut). Par exemple,

```
t.test(IB[,1])
```

Comparer les intervalles de confiance du groupe de référence et du groupe traité.
Conclure.