

## Cours d'analyse informatique des données

### Séance de laboratoire 4

#### Commandes utiles

`plot` : graphe d'un signal (points reliés).

`subplot` : permet d'avoir plusieurs graphes sur la même figure.

`mean` : estimation de la moyenne.

`var` : estimation de la variance.

`std` : estimation de la déviation standard.

`median` : estimation de la médiane.

#### Commandes supplémentaires

`somme_uniformes`: génération réalisation de sommes de VA uniformes

`somme_uniformes_sq`: génération réalisation de sommes de VA uniformes au carré

`hist_comp_gaussien`: estimation de densité par histogramme et représentation gaussienne.

`gaussienne_perturbee` : génération VA gaussienne contaminée

`stat_cardio` : statistiques données cardiovasculaires

#### Expérience 1 : vérifions le théorème central limite

Ce théorème démontre qu'une somme de variables indépendantes non gaussiennes tend à devenir gaussienne quand le nombre de variables sommées augmente.

1. La routine `somme_uniformes` génère 1000 réalisations d'un somme de  $K$  VA de densité uniforme entre 0 et 1. La routine `hist_comp_gaussien` met sur un graphe la densité de probabilité estimée par histogramme et la densité gaussienne correspondant à la moyenne et variance estimées sur les réalisations. En commençant avec  $K = 1$ , vérifiez avec `hist_comp_gaussien` que les réalisations obtenues avec `somme_uniformes` ont une densité estimée de plus en plus proche de la gaussienne. Avec `subplot`, représentez 4 exemples représentatifs de cette progression.

2. Refaites la même expérience avec `somme_uniformes_sq`. Notez que la progression vers la gaussienne est plus lente, car la densité initiale est plus éloignée d'une gaussienne.

#### Expérience 2: médiane parfois mieux que moyenne

La routine `gaussienne_perturbee` génère 1000 réalisations d'une VA gaussienne de moyenne 0 et de variance 1, contaminée par un bruit uniforme positif, la probabilité de contamination par réalisation étant donnée par le paramètre  $q$ . Générez des réalisations pour  $q = 0, 0.05, 0.1, \text{ et } 0.2$ . Visualisez pour chaque valeur la densité avec `hist_comp_gaussien` et estimez la moyenne et la médiane. Quelle mesure est la moins perturbée?

#### Expérience 3: l'alcool influence les statistiques

Le fichier `heart_1.dat` contient 3 colonnes, correspondant à 3 signaux acquis simultanément à une fréquence d'échantillonnage de 4 Hz. La première colonne

correspond aux intervalles RR (intervalles entre battements cardiaques, en millisecondes), la seconde est la pression artérielle (en mm de mercure) et la troisième le volume pulmonaire instantané (respiration, unités arbitraire, en fait même signal que dans le labo 1).

Une fois le fichier chargé avec `load`, récupérez les intervalles RR avec :

```
>> RR = heart_1(:,1);
```

1. Avec `hist_comp_gaussien` vérifiez que la densité des intervalles RR est approximativement gaussienne. Appliquez la routine `stat_cardio`, qui calcule moyennes, variances et covariances normalisées à `heart_1(:,1)`. Est-il normal que la covariance entre intervalles RR et respiration soit négative ?

2. Refaites ceci avec les données de **heart\_2.dat** (mêmes colonnes, même fréquence d'échantillonnage). Il s'agit du même sujet, mais 20 minutes après une absorption massive d'alcool qui suit le premier enregistrement. Quels sont les changements marquants ?