



Tests d'hypothèse : détection d'un son



Crédit : Jurassic Park

Formateur : Raphael Bacher

Objectifs pédagogiques

Utiliser les outils statistiques standards permettant de réaliser un test d'hypothèse, et réaliser ce test.

Un test d'hypothèse statistique est la vérification d'une hypothèse portant sur une variable aléatoire, sur la base d'observations (mesures) de ladite variable. La conclusion du test sera le rejet ou le non-rejet de l'hypothèse. Dès lors qu'on tire une conclusion d'un jeu de donnée, il est nécessaire d'associer un critère de significativité statistique à cette conclusion.

Il existe une multitude de tests d'hypothèse en statistique. On peut par exemple comparer 2 échantillons afin de déterminer s'ils sont statistiquement significativement différents. On peut aussi comparer un jeu de données à un modèle statistique censé décrire le comportement d'une variable aléatoire.

L'objectif de ce module est triple :

- Que vous vous formiez à la construction de lois de probabilités (probability distributions).
- Que vous appréhendiez via un exemple les caractéristiques systématiquement présentes dans un test d'hypothèse. Et ce de manière à ce que vous soyez en mesure de vous approprier dans le futur les tests d'hypothèse que vous pourrez avoir à réaliser.
- Si vous avez le temps, de méditer plus en profondeur sur ce qu'on peut dire d'un test d'hypothèse et ce que l'on ne peut/doit pas.

Modalités

- Durée du projet : 2.5 jours.
- Ce projet sera réalisé en autonomie
- Mais l'entraide est très fortement encouragée !
- Vous produirez un jupyter notebook contenant la résolution/réponse aux questions du sujet.

Démarche pédagogique

Pour les 3 journées de cours consacrés à ce projet, nous allons nous intéresser à la détection de signal sonore. Cela nous amènera en première itération à réaliser une étude sur un bruit. En seconde itération, nous réaliserons une étude sur un son. Puis en troisième et dernière itération nous nous intéresserons à la détection d'un signal sonore présent au milieu du bruit. Cela constituera votre test d'hypothèse : pour réaliser une détection, nous allons devoir faire une hypothèse qui nous permettra de décider si oui ou non, un signal sonore est détecté.

Vous aborderez les notions suivantes :

- Lois de probabilités en statistique (continues et discrètes). Probability distributions.
- Eléments de probabilités
- Tests d'hypothèses (choix d'une hypothèse nulle, d'une hypothèse alternative, critère de significativité et p-value)

Si cela remonte à loin, vous aurez besoin de vous rafraîchir la mémoire sur certaines notions :

- Les bases du calcul différentiel et intégral (Qu'est ce qu'une dérivée ? Une intégrale ? etc). Vous n'aurez pas besoin de calculer des intégrales analytiquement mais il sera important de comprendre ces concepts pour avancer
- Les bases de probabilités.

Sujet : Depuis quelque temps, bien que vivant dans un endroit très calme la nuit, vous vous demandez si des sons ne vous dérangent pas pendant votre sommeil. Bien décidé à avoir le coeur net et plutôt que de vous demander à chaque réveil si celui-ci est dû à son quelconque ou juste à une mauvaise qualité de sommeil, vous décidez de réaliser un enregistrement sonore d'une nuit, et de créer un test qui permettra de détecter les sons qui se seront produits durant cette nuit.

Tous les enregistrements sont disponibles avec le sujet, chacun en double :

- un au format ".wav" (ça serait dommage de ne pas utiliser vos oreilles alors que vous travaillez sur des signaux sonores)
- l'autre et en format .txt.

Chaque enregistrement est **échantillonné à 48 kHz**.

Pour éviter de perdre du temps, la variable "signal sonore" contenue dans chaque fichier .txt est déjà normalisée, pour avoir des amplitudes de signaux n'excédant pas quelques unités. Vous n'aurez donc pas besoin de modifier vos données pour travailler dessus.

Compétences

- Construire les fonctions de distribution et répartition d'une variable aléatoire continue
- Construire les fonctions de distribution et répartition d'une variable aléatoire discrète
- Construire mener à bien et interpréter un test d'hypothèse

Consignes

- Voir notebook

Livrables

- Exercices terminés dans le jupyter notebook