

Biostatistiques appliquées

Fondamentaux et tests statistiques usuels avec R (3 jours – 21 h)

A qui s'adresse cette formation ?

Cette formation s'adresse à des personnes souhaitant mieux comprendre et utiliser les analyses statistiques : interpréter les résultats, évaluer la pertinence des méthodes employées et réaliser des analyses usuelles sous R, de manière rigoureuse.

Elle est particulièrement adaptée aux :

- professionnel(le)s de la santé, de l'environnement ou de la recherche,
- ingénieur(e)s, chargé(e)s d'études, doctorant(e)s ou post-doctorant(e)s,
- personnes utilisant R (ou souhaitant l'utiliser) pour analyser leurs données,
- personnes amenées à lire, produire ou interpréter des analyses statistiques.

Cette formation pose les bases du raisonnement statistique et de la mise en œuvre des tests usuels sous R.

Elle pourra être complétée par une formation de niveau avancé dédiée à la modélisation statistique, dont l'ouverture est prévue à partir de septembre 2026.

Ce que cette formation n'est pas

Cette formation n'a pas vocation à :

- enseigner l'ensemble des biostatistiques en 3 jours,
- se substituer à un cursus académique en statistiques,
- fournir des "recettes toutes faites" sans réflexion méthodologique.

Elle vise au contraire à poser des bases solides, à éviter les erreurs d'interprétation, et à développer une approche rigoureuse et raisonnée des tests statistiques usuels.

Prérequis

Avoir déjà été exposé(e) aux notions de base en statistiques ou en biostatistiques (types de variables, tests statistiques usuels, p-value)

Connaître les bases du logiciel R : manipulation d'objets simples, importation de données, exécution de scripts et compréhension d'un code R existant.

Objectifs et compétences visées

Cette formation vise à permettre aux participant(e)s de raisonner correctement autour des tests statistiques usuels, d'en comprendre les principes, et d'en interpréter les résultats de manière rigoureuse, tout en sachant les mettre en œuvre sous R.

À l'issue de la formation, les participant(e)s seront capables de :

- identifier le type de variables et formuler correctement une question statistique,
- choisir un test statistique usuel adapté au contexte d'étude et aux données disponibles,
- comprendre et vérifier les hypothèses associées aux tests statistiques,
- appliquer correctement les tests statistiques sous R (choix des fonctions, paramétrage, lecture des sorties)
-
- interpréter les résultats statistiques (statistiques de test, p-values, intervalles de confiance) sans contresens,
- analyser de façon critique les résultats obtenus et en apprécier les limites,
- présenter et commenter des résultats statistiques de manière claire, structurée et rigoureuse,
- gagner en autonomie et en fiabilité dans la réalisation et l'interprétation d'analyses statistiques sous R.
-

Ces compétences permettent de sécuriser les analyses statistiques, de mieux justifier les choix méthodologiques et de produire des résultats interprétables et exploitables dans un cadre professionnel.

Programme de la formation

Module 1 – Décrire et explorer ses données

Objectif : comprendre la structure des données et produire des descriptions pertinentes.

Types de données et de variables

Indicateurs descriptifs (position, dispersion, distribution)

Tableaux de comptage, de proportions et tableaux croisés

Construction de tableaux descriptifs simples et complexes

Graphiques descriptifs et principes de data visualisation

Mise en œuvre sous R

Module 2 – Estimer et quantifier l'incertitude

Objectif : comprendre ce que l'on estime à partir d'un échantillon et comment interpréter la précision.

Notions de population, échantillon et paramètres

Estimations ponctuelles

Intervalles de confiance et notion de précision

Méthodes de construction et interprétation des intervalles de confiance

Calcul de taille d'échantillon pour une précision donnée

Applications pratiques sous R

Module 3 – Comprendre la logique des tests statistiques

Objectif : savoir pourquoi et comment un test est utilisé, au-delà de la formule.

Cadre général des tests statistiques

Hypothèses nulle et alternative

Statistique de test et loi théorique

Règle de décision, risque alpha et p-value

Formulation correcte des conclusions

Conditions d'application des tests

Tests paramétriques et non paramétriques

Indépendance, appariement

Notions de tests d'équivalence et de non-infériorité

Panorama des principaux tests de supériorité

Module 4 – Comparer des moyennes

Objectif : analyser et interpréter des comparaisons de groupes quantitatifs.

Comparaison d'une moyenne à une valeur de référence

Comparaison de deux moyennes indépendantes

Comparaison de deux moyennes appariées

Comparaison de plusieurs moyennes : ANOVA à un facteur

Mise en œuvre et interprétation sous R

Module 5 – Comparer des pourcentages

Objectif : analyser des données qualitatives et des fréquences.

Comparaison d'un pourcentage à une valeur de référence

Comparaison de deux pourcentages indépendants

Comparaison de deux pourcentages appariés

Choix du test et interprétation des résultats

Applications sous R

Module 6 – Étudier le lien entre deux variables

Objectif : analyser des relations entre variables selon leur nature.

Lien entre deux variables quantitatives : corrélation, régression linéaire simple, tests associés

Lien entre deux variables qualitatives : Test du Chi-2

Interprétation des résultats

Mise en pratique sous R

Module 7 – Puissance statistique et taille d'échantillon

Objectif : comprendre les enjeux de puissance et dimensionner correctement une étude.

Risque beta et puissance statistique

Notion d'effect size

Calculs de puissance et de taille d'échantillon

Cas des comparaisons de moyennes et de proportions

Applications sous R

Module 8 - Aller plus loin dans l'analyse des données

Objectif : découvrir les modèles multivariés et les méthodes exploratoires.

Intérêt des modèles de régression

Présentation des modèles de régression linéaire

Introduction aux modèles linéaires généralisés

Études de cas pratiques

Introduction à l'Analyse en Composantes Principales (ACP)

Cas pratiques et interprétation


Utilisation raisonnée de l'intelligence artificielle


Tout au long de la formation, des outils d'intelligence artificielle sont utilisés comme supports pédagogiques afin de clarifier des concepts statistiques, questionner le choix des méthodes, vérifier des raisonnements et des interprétations, et accompagner la mise en œuvre des analyses sous R.

L'IA est mobilisée comme outil d'assistance au raisonnement statistique, dans une démarche critique, encadrée et complémentaire des apprentissages

Modalités pratiques

 **Durée** : 3 jours (21 heures)

 **Format** : distanciel (visioconférence), présentiel possible selon les besoins


 **Effectif** : Groupes à taille limitée (environ 5), favorisant les échanges et les questions


 **Pédagogie et organisation** :

Alternance d'apports théoriques et de mises en pratique

Supports et scripts commentés remis aux participant(e)s

Temps dédiés aux questions et à l'interprétation

 **Logiciels** : R et R Studio, Git, G*Power (calcul de puissance / nombre de sujets nécessaires)

 **Évaluation des acquis** : Exercices pratiques, mises en situation et échanges oraux tout au long de la formation.

 **Tarif**

1 500 € HT par participant(e) pour la formation de 3 jours (21 h) - Supports pédagogiques inclus.

Les tarifs en intra-entreprise sont établis sur devis, en fonction du format et des modalités de la formation.

Organisme exonéré de TVA en vertu de l'article 261-4-4°a du CGI.

✓ **Suivi et assiduité** : Émargement par demi-journée.

 **Adaptation et accessibilité** :

Les exemples peuvent être adaptés au contexte des participants en intra-entreprise.

Toute situation de handicap ou besoin spécifique peut être signalé en amont afin d'étudier les adaptations possibles

Informations de contact

Organisme de formation : **DellaDATA**

NDA : 93040100204 auprès du préfet de région PACA

SIRET : 90838795400013

Organisme exonéré de TVA en vertu de l'article 261-4-4°a du CGI.

Adresse : 1087 CH de Ste Roustagne - 04100 Manosque

Formatrice et interlocutrice : Claire Della Vedova (claire@delladata.fr / 06.64.98.21.42)