Formules de probabilités et statistique

Probabilités et statistique pour la biologie (STAT1)

Jacques van Helden

2018-05-10

Table of Contents

# Combinatoire

|  |  |  |
| --- | --- | --- |
| Nom | Conditions | Formule |
| Permutations (factorielle) |  |  |
| Arrangements | Sans remise, ordonné |  |
| Combinaison (*choose*, *coefficient binomial*) | Sans remise, sans ordre |  |
|  |  |  |

# Concepts de probabilité

|  |  |  |
| --- | --- | --- |
| Description | Conditions | Formule |
| Définition fréquentielle de la probabilité |  |  |
|  |  |  |
| Probabilité de non-réalisation |  |  |
|  |  |  |
| Probabilités conditionnelles |  |  |
|  |  |  |
|  |  |  |
| Probabilité de ou | En général |  |
|  |  |  |
| Probabilité de ou | Evénements mutuellement exclusifs |  |
|  |  |  |
| Probabilité de et | En général |  |
|  |  |  |
| Probabilité de et | Evénements indépendants |  |
|  |  |  |
| Règle de Bayes |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# Distributions de probabilité discrètes

## Géométrique

* Conditions : nombre d’échecs avant le premier succès dans un schéma de Bernoulli
* Densité :
* Répartition :
* Moyenne :
* Variance :

## Binomiale

* Conditions : Nombre de succès au cours d’une série d’essais indépendants avec probabilité constante de succès (Schéma de Bernoulli)
* Densité :
* Répartition :
* Moyenne :
* Variance :
* Rapport moyenne/variance:

## Poisson

* Conditions : nombre de succès observés au cours d’un intervalle de temps, en fonction du nombre attendu ()
* Application : approximation de la binomiale quand et faible ()
* Densité :
* Répartition :
* Moyenne :
* Variance :
* Rapport moyenne/variance:

## Hypergéométrique

* Conditions : Tirage non ordonné, sans remise dans un ensemble fini avec deux catégories.
* Exemple-type: urne avec boules de deux couleurs
* Densité :
* Répartition :
* Moyenne :
* Variance :

# Echantillonnage et estimation

* Les symboles grecs (, ) correspondent aux statistiques de population, les symboles romains (, ) aux statistiques d’échantillon.
* L’accent circonflexe () indique les estimateurs de paramètres de population calculés à partir de paramètres d’échantillons.

|  |  |
| --- | --- |
| Symbole | Description |
|  | Taille (nombre d’individus) de la population. |
|  |  |
|  | Moyenne de la population. |
|  |  |
|  | Variance de la population |
|  |  |
|  | Écart-type de la population |
|  |  |
|  | Effectif (nombre d’individus) de l’échantillon. |
|  |  |
|  | Moyenne d’échantillon. |
|  |  |
|  | Variance de l’échantillon |
|  |  |
|  | Écart-type de l’échantillon |
|  |  |
|  | Estimateur non-biaisé de la variance de la population. |
|  |  |
|  | Estimateur non-biaisé de l’écart-type de la population. |
|  |  |
|  | Erreur standard: écart-type attendu sur la moyenne d’échantillon. |
|  |  |
|  | Intervalle de confiance autour de la moyenne. |
|  |  |

# Test de comparaison de moyennes

|  |  |
| --- | --- |
| Symbole | Description |
|  | Moyennes respectives des populations 1 et 2. |
|  |  |
|  | Écarts-types respectifs des populations 1 et 2. |
|  |  |
| , | Effectifs (nombre d’individus) des échantillons prélevés sur les populations 1 et 2. |
|  |  |
|  | Formule générale de la moyenne d’échantillon |
|  |  |
|  | Moyennes d’échantillons. |
|  |  |
|  | Différence entre les moyennes des populations. |
|  |  |
|  | = **Taille d’effet**: dans un test de comparaison de moyennes, il s’agit de la différence entre les moyennes d’échantillons, utilisée comme estimateur de . |
|  |  |
|  | Formule générale de la variance d’échantillon |
|  |  |
|  | Variances mesurées sur les échantillons. |
|  |  |
|  | Écart-type groupé (*pooled standard deviation*), utilisé comme estimateur de l’écart-type commun des deux populations, en supposant leurs variances égales (hypothèse de travail d’homoscédasticité). |
|  |  |
|  | Erreur standard sur la différence entre moyennes, en supposant que les populations ont la même variance (test de Student). |
|  |  |
|  | Statistique de Student, d.l. |
|  |  |
|  | Statistique de Welch, d.l. |
|  |  |