

TEST D'HYPOTHÈSES (I1, I3)

(09 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

La progression de l'activité scientifique (cf **science**, **étude scientifique**) nécessite, de façon générale, des bases assurées : cette démarche suit la **méthode des pierres de gué**. A chaque période de son évolution, diverses **questions** doivent, dans la mesure du possible, recevoir des **réponses**. La réponse (issue) à une question donnée est donc, a priori, soit positive, soit négative, soit dubitative (doute), et elle conditionne la direction de l'activité de la période suivante.

Dans tous les cas, il existe des risques (probabilités), plus ou moins important(e)s, de jugements erronés. En effet, il est très rare que les réponses aux questions soient absolues, puisque le statisticien raisonne toujours implicitement en **univers aléatoire** (cf **contexte statistique**), et que des déterminismes de type mathématique ne s'y rencontrent guère, quel que soit le **domaine de connaissance**.

En **Statistique**, l'investigation d'un **phénomène** relevant d'un domaine donné fait, en règle générale, l'objet d'une **modélisation** (théorisation) qui doit aussi permettre la **validation des réponses aux questions**. Les méthodes de test d'hypothèses ont pour finalité cette validation : cette dernière est donc assortie d'un degré de confiance, variable en fonction des observations mais aussi de la **spécification** retenue.

(i) Un **test d'hypothèses statistique(s)**, ou **test statistique**, ou le plus souvent simplement **test**, est une **procédure statistique** destinée à indiquer au **statisticien** s'il peut tenir une **hypothèse statistique** pour « vraie » : autrement dit, il s'agit d'une procédure dont l'objet est de fournir un « degré de plausibilité », ou « degré de vraisemblance », à l'hypothèse en question (celle privilégiée), sachant qu'il existe, le plus souvent, des hypothèses alternatives (ou hypothèses concurrentes).

Un test est donc un instrument aidant à répondre à diverses questions, et notamment à arbitrer entre une théorie donnée et d'autres théories (théories alternatives antérieures ou contemporaines, théories nouvelles) (cf **théorie des tests**).

(ii) Soit $(\Omega, \mathcal{F}, \mathcal{P})$ un **modèle statistique** fondamental, $(\mathcal{X}, \mathcal{B})$ un **espace d'observation** et $(\mathcal{X}, \mathcal{B}, \mathcal{P}^{\mathcal{X}})$ le **modèle image** du précédent par une **va** (eg **échantillon**) $X : \Omega \mapsto \mathcal{X}$.

(a) on appelle **fonction de test**, ou simplement **test**, une **application mesurable** (ie une **statistique**) $\delta : \mathcal{X} \mapsto \{0, 1\} = \mathbb{N}_1$.

Cette notion est à distinguer de celle de **statistique de test** : une statistique de test permet de définir des **régions critiques** associées à divers risques d'espèce, ie principalement à deux types de risque : le **risque de première espèce** et le **risque de seconde espèce**.

La probabilité de chacune des deux issues est représentable dans une **table de décision** 2 x 2 (cf ci-après) : la probabilité de décider d_a (ie l'état « vrai » de la Nature est H_a) alors que l'état « vrai » de la Nature est H_0 est généralement notée α , tandis que la probabilité (« duale ») de décider d_0 (ie l'état « vrai » de la Nature est H_0) alors que l'état de la Nature est H_a est notée β . Le complément à l'unité de cette dernière, ie $\eta = 1 - \beta$, est appelé **puissance** du test considéré ;

		états de la Nature	
		H_0	H_a
décisions du statisticien	d_0	$1 - \alpha$	β
	d_a	α	$1 - \beta = \eta$

α = risque de première espèce
 β = risque de seconde espèce
 η = puissance

(b) la notion plus générale de **test mixte** (ou test aléatoire) constitue un outil de base dans les problèmes de tests.

(iii) On recherche souvent un test possédant la propriété d'être « sensible » aux **hypothèses alternatives** H_a : un test doit aboutir à « abandonner » une **hypothèse de base** H_0 (hypothèse en général privilégiée) si les **observations** impliquent une telle « évidence ».

Il existe, en effet, des tests trop « conservateurs », ie qui rejettent peu souvent l'hypothèse testée H_0 ; mais il existe aussi des tests trop « novateurs », lesquels rejettent trop souvent H_0 .

Dans la **théorie de NEYMAN - PEARSON** (cf **test de NEYMAN-PEARSON**), on fixe généralement le niveau α_0 du **risque de première espèce** α ou, alternativement, la **loi a priori** de α . Lorsque le test de H_0 est réalisé et que H_0 est acceptée pour vraie, on peut aussi chercher à déterminer le plus petit seuil $\alpha_{\sim} < \alpha_0$ (appelé **niveau de confiance**) tq H_0 est encore acceptée.