Chapitre III

Tests statistiques

1 Introduction

Depuis quelques décennies, nous assistons à une "arrivée en force" des méthodes statistiques dans le domaine réglementaire, lequel conduit à la prise de décision : on a ou on n'a pas le droit de … En particulier, l'augmentation des échanges commerciaux et des liens économiques entre les pays s'accompagne d'accords destinés à fixer des règles communes; la statistique inférentielle trouve là un immense champ d'application. Cela se traduit par des réglementations définissant dans chaque cas particulier une procédure destinée à préciser sans ambiguïté :

- comment un ou plusieurs échantillons doivent être prélevés dans la population étudiée;
- quelles mesures doivent être effectuées sur ce ou ces échantillons;
- quelle décision doit être prise à propos de l'ensemble de la population étudiée, suivant les résultats obtenus sur le ou les échantillons.

Une telle procédure s'appelle en statistique un *test de validité d'hypothèse.*

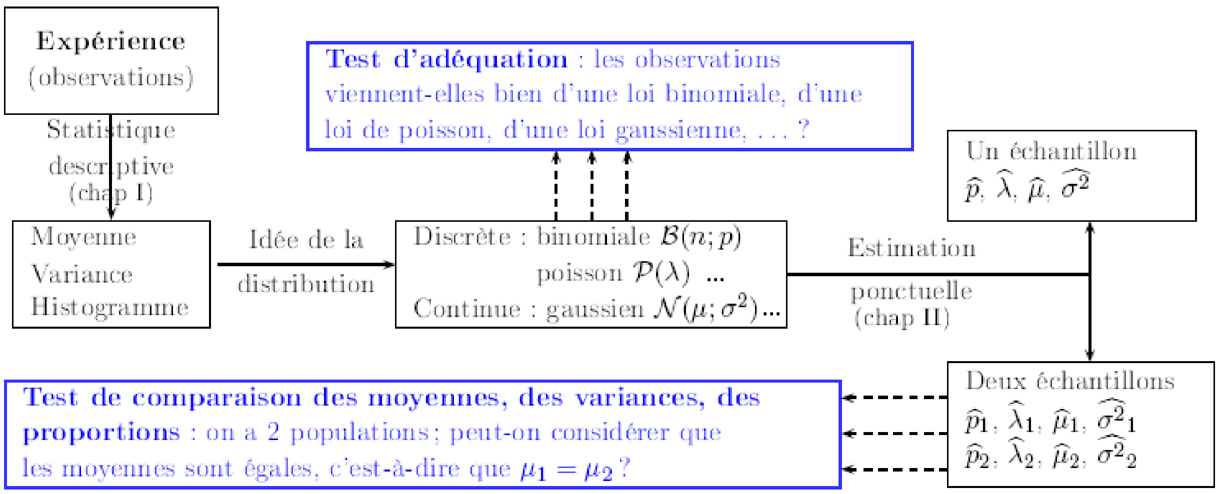


FIG. III.1 - Diagramme récapitulatif : pourquoi les tests ?

Nous allons étudier plusieurs types de tests : test d'adéquation à une loi théorique (qui permet de valider l'ajustement d'une distribution expérimentale issue d'un échantillon à une loi théorique), test d'indépendance de deux caractères, tests de comparaison ...

La construction d'un test, quel qu'il soit, passe par deux étapes :

1. On formule une hypothèse H_0 appelée *hypothèse nulle*, et il s'agit de décider si on rejette cette hypothèse par opposition à une contre-hypothèse H_1 appelée *hypothèse alternative*.

2. On applique une *règle de décision* qui va nous permettre de choisir entre la première hypothèse et la seconde. Elle repose sur une *variable de décision* (qui est une statistique) dont on connait la loi si (H_0) est vraie.

Nous pouvons alors résumer les différentes probabilités suivant la réalité et notre décision, dans le tableau ci-dessous :

		Réalité (ou état de la nature)	
		H_0 vraie	H_1 vraie
Décision	considérer H_0 vraie	$1-\alpha$	β
retenue	considérer H_1 vraie	α	$1-\beta$

TAB. III.1 - Probabilités intervenant lors d'une prise de décision en environnement incertain.

Définition 23 La quantité α représente le pourcentage de cas où H_0 est vraie et on la rejette. On l'appelle *risque de première espèce* ou encore *niveau* ou *seuil de signification*.
(En traitement du signal, α s'appelle probabilité de fausse alarme.)

La quantité β représente le pourcentage de cas où H_1 est vraie et on la rejette.

On l'appelle *risque de seconde espèce*. La quantité $1 - \beta$ s'appelle la *puissance du test*.

Lorsque l'on est en présence d'un tel test, ce que l'on cherche à faire est de minimiser les erreurs, c'est à dire les valeurs α et β. La règle de décision est très importante puisqu'elle va induire le calcul de α et de β. Pour minimiser ces deux valeurs, il faut donc jouer sur cette règle de décision. Dans la plupart des cas, nous ne pouvons pas jouer sur les deux tableaux, c'est à dire minimiser à la fois α et β. Nous allons étudier le cas où l'on donne α.

2 Tests basés sur le Khi-2

2.1 Test d'adéquation

On étudie un phénomène aléatoire représenté par une V.A. X.

On considère un n-échantillon $(X_1, \ldots, X_n)$ que l'on analyse selon les méthodes de statistique descriptive.
Cela permet de choisir parmi les lois de probabilités usuelles celle qui semble être la plus proche de la distribution expérimentale induite par l'échantillon.

Remarque 20 on peut être amené à ce stade à estimer certains paramètres de la loi choisie (ou donnée) pour modéliser le phénomène aléatoire (comme le paramètre λ d'une loi de Poisson).

A partir de ces données, on veut tester l'hypothèse suivante :

H_0 : "X suit une certaine loi de probabilité L"

Cela revient à tester l'hypothèse : " $X_1, \ldots, X_n$ suivent la loi de probabilité L"

Soit $(x_1,\ldots, x_n)$ une réalisation du n- échantillon.

On note $m_1,\ldots, m_l$ les valeurs distinctes prises par $x_1,\ldots, x_n$ et y_i le nombre de fois où m_i apparaît dans $(x_1,\ldots, x_n)$ (il s'agit de *l'effectif observé* pour la modalité m_i).

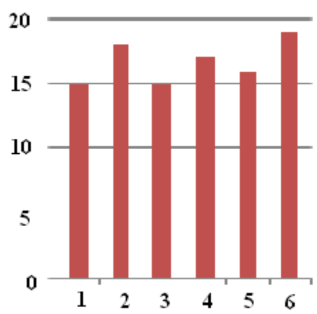
Exemple 25 On effectue 100 lancers d'un dé équilibré (cas d'équiprobabilité) à 6 faces. On obtient le tableau suivant :

Face	1	2	3	4	5	6
Fréquence	15	18	15	17	16	19

TAB. III.2 – Résultats de l'expérience du jet du dé à 6 faces

La question est de savoir si le dé est vraiment équilibré ou non.

Le diagramme en bâtons correspondant à notre distribution expérimentale est le suivant :



Le diagramme ressemble bien à une loi uniforme discrète sur {1, 2, 3, 4, 5, 6}.

La loi d'équiprobabilité qui s'applique en probabilités si le dé est équilibré donne le tableau suivant :

Face	1	2	3	4	5	6
Fréquence	$100\times\frac{1}{6}$	$100\times\frac{1}{6}$	$100\times\frac{1}{6}$	$100\times\frac{1}{6}$	$100\times\frac{1}{6}$	$100\times\frac{1}{6}$

TAB. III.3 - Résultats théoriques du jet d'un dé à 6 faces

A ce niveau, on peut faire plusieurs remarques :

- Il n'est évidemment pas possible de faire 16,67 fois le chiffre 1. Il faut une valeur entière! Il y a donc dans la plupart des cas, une différence entre la fréquence théorique et la fréquence expérimentale (même si le dé est parfait).

- La différence entre la fréquence théorique (16,67) et la fréquence expérimentale (15 à 19) ne semble pas "significative". Il semble alors normal de dire que le dé est équilibré. Toutefois, et pour être plus rigoureux, nous allons "mesurer" cet écart. Cette "mesure" (qui donnera la variable de décision) nous permettra de dire avec une certaine confiance (95% ou 99%) si la différence est "significative".

Dans cet exemple, on a :

- $m_1 = 1;\ m_2 = 2;\ m_3 = 3;\ m_4 = 4;\ m_5 = 5;\ m_6 = 6,$
- $y_1 = 15;\ y_2 = 18;\ y_3 = 15;\ y_4 = 17;\ y_5 = 16;\ y_6 = 19.$

Il faut alors comparer les fréquences absolues théoriques pour m_i ($n\times \mathrm{P}(X = m_i)$) avec les fréquences expérimentales y_i.

On mesure l'écart entre la distribution théorique et la distribution expérimentale par la quantité suivante :

$$\boxed{t_n = \sum_{i=1}^{l} \frac{\left[y_i - n\mathrm{IP}(X = m_i)\right]^2}{n\mathrm{IP}(X = m_i)}}$$

Plus t_n est grand, plus la divergence entre la distribution théorique et la distribution expérimentale est grande. Il est donc normal de rejeter H_0 lorsque t_n dépasse une quantité c_α qui dépend du risque de première espèce.

La règle de décision convenable est donc:

On rejette H_0 ssi $t_n > c_\alpha$.

Le prochain résultat va nous fournir la possibilité de calculer explicitement c_α.

Théorème 7 Soit Y_i la V.A. donnant le nombre d'observations égales à m_i dans le n-échantillon. Si on suppose que $X_1,\ldots, X_n$ ont même loi que X (c'est à dire que l'hypothèse H_0 est vraie), la variable de décision

$$T_n = \sum_{i=1}^{l} \frac{\left[Y_i - n\mathrm{IP}(X = m_i)\right]^2}{n\mathrm{IP}(X = m_i)}$$

suit un loi du Khi-2 à (l - 1) degrés de liberté (d.d.l.).

Conséquence 1 On a la règle de décision suivante :

On rejette H_0 ssi $t_n > \chi^2_{l-1;1-\alpha}$

où $\chi^2_{l-1;1-\alpha}$ est le fractile d'ordre 1 - α d'une loi du Khi-2 à (l - 1) d.d.l., c'est-à-dire :

$\mathrm{IP}\left(\chi^2_{l-1} < \chi^2_{l-1;1-\alpha}\right) = 1-\alpha$ où χ^2_{l-1} suit une loi du Khi-2 à (l - 1) d.d.l.

(Pour les valeurs de $\chi^2_{l-1;1-\alpha}$ cf annexe 3).

Exemple 26 Pour le dé, on a : t_n = 0,799 et $\chi^2_{5;95\%}$ = 11,07. Comme $t_n < \chi^2_{5;95\%}$, on accepte H_0. On considère donc que le dé est équilibré.

Remarque 21 Si la loi théorique dépend de paramètres inconnus (comme par exemple, si on a estimé le paramètre λ de la loi de Poisson), la règle de décision doit prendre en compte les éventuelles estimations. On a alors la règle de décision suivante :

On rejette H_0 ssi $t_n > \chi^2_{l-p-1;1-\alpha}$ où p est le nombre de paramètres estimés.

2.2 Test d'indépendance de deux variables qualitatives

2.2.1 Table de contingence

Soient, dans une même population, deux caractères qualitatifs (appelés *facteurs*) :

- le caractère L ayant l modalités,
- le caractère C ayant c modalités.

On prélève au hasard n individus et on note x_{ij} le nombre d'observations de la cellule (L_i; C_j), c'est à dire le nombre d'individus possédant la $i^{ème}$ modalité de L et la $j^{ème}$ modalité de C, avec $1 \leq i \leq l$ et $1 \leq j \leq c$.

On dispose alors d'une table de contingence dans laquelle chacun des n individus doit se retrouver dans une seule cellule.

$L \setminus C$	C_1	$\cdots$	C_j	$\cdots$	C_c	Total
L_1	x_{11}	$\cdots$	x_{1j}	$\cdots$	x_{1c}	$x_{1\bullet}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
L_i	x_{i1}	$\cdots$	x_{ij}	$\cdots$	x_{ic}	$x_{i\bullet}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
L_l	x_{l1}	$\cdots$	x_{lj}	$\cdots$	x_{lc}	$x_{l\bullet}$
Total	$x_{\bullet 1}$	$\cdots$	$x_{\bullet j}$	$\cdots$	$x_{\bullet c}$	n

TAB. III.4 – Tableau de contingence

On calcule les effectifs marginaux par :

$$x_{i\bullet} = \sum_{j=1}^{c} x_{ij} \quad \text{et} \quad x_{\bullet j} = \sum_{i=1}^{l} x_{ij}$$

De plus, on a :

$$\sum_{i=1}^{l}\sum_{j=1}^{c} x_{ij} = \sum_{i=1}^{l} x_{i\bullet} = \sum_{j=1}^{c} x_{\bullet j} = n$$

2.2.2 Règle de décision

Les hypothèses du test sont :

$$\begin{cases} H_0 : \text{les deux caractères sont indépendants} \\ H_1 : \text{ non } H_0 \end{cases}$$

On note p_{ij} la probabilité pour un individu d'appartenir à la cellule (L_i; C_j) pour $1 \leq i \leq l$ et $1 \leq j \leq c$. On déduit les probabilités marginales $p_{1\bullet}, ..., p_{i\bullet}, ..., p_{l\bullet}$ pour le caractère L et $p_{\bullet 1}, ..., p_{\bullet j}, ..., p_{\bullet c}$ pour le caractère C.

On sait que L et C sont indépendantes en probabilité si $p_{ij} = p_{i\bullet}.p_{\bullet j}$ pour tout (i ; j).

On admet que $p_{i\bullet}$ et $p_{\bullet j}$ sont estimés respectivement par $\frac{x_{i\bullet}}{n}$ et $\frac{x_{\bullet j}}{n}$.

Il faut alors estimer (l + c – 2) paramètres car $p_{l\bullet}$ et $p_{\bullet c}$ se déterminent par les relations $\sum_{j=1}^{c} p_{\bullet j} = \sum_{i=1}^{l} p_{i\bullet} = 1$.

On note $n_{ij} = \frac{x_{i\bullet}\, x_{\bullet j}}{n}$, appelé l'*effectif théorique*.

Quand H_0 est vraie, on peut estimer p_{ij} par $\frac{n_{ij}}{n^2}$.

On mesure l'écart entre les effectifs théoriques et expérimentaux par la quantité suivante :

$$\boxed{t_n = \sum_{i=1}^{l}\sum_{j=1}^{c} \frac{\left(x_{ij} - n_{ij}\right)^2}{n_{ij}}}$$

Théorème 8 Soient X_{ij} et N_{ij} les V.A. dont les réalisations sont respectivement x_{ij} et n_{ij}.
Si on suppose que l'hypothèse H_0 est vraie, la variable de décision

$$T_n = \sum_{i=1}^{l}\sum_{j=1}^{c} \frac{\left(X_{ij} - N_{ij}\right)^2}{N_{ij}}$$

suit un loi du Khi-2 à (l – 1) (c – 1) degrés de liberté (d.d.l.).

En effet, le degré de liberté de la loi limite est égal à : $lc - (l + c - 2) - 1 = (l - 1)(c - 1)$

Conséquence 2 On a la règle de décision suivante :

$$\boxed{\text{On rejette } H_0 \quad \text{ssi} \quad t_n > \chi^2_{(l-1)(c-1);1-\alpha}}$$

2.3 Test d'homogénéité d'une V.A.

2.3.1 Table de contingence

On considère l populations $P_1, \ldots, P_l$ sur lesquelles on étudie une V.A. X prenant c modalités $m_1, \ldots, m_c$.

Définition 24 On dit que les populations sont *homogènes* si la distribution de X est la même dans les l populations.

On prélève au hasard n individus répartis dans l'ensemble des populations et on note x_{ij} le nombre d'individus de la population P_i possédant la modalité m_j avec $1 \le i \le l$ et $1 \le j \le c$.
On dispose alors d'une table de contingence dans laquelle chacun des n individus doit se retrouver dans une seule cellule :

$P \backslash m$	m_1	$\cdots$	m_j	$\cdots$	m_c	Taille des échantillons
P_1	x_{11}	$\cdots$	x_{1j}	$\cdots$	x_{1c}	n_1
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
P_i	x_{i1}	$\cdots$	x_{ij}	$\cdots$	x_{ic}	n_i
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
P_l	x_{l1}	$\cdots$	x_{lj}	$\cdots$	x_{lc}	n_l
Effectifs marginaux	$x_{\bullet 1}$	$\cdots$	$x_{\bullet j}$	$\cdots$	$x_{\bullet c}$	n

TAB. III.5 – Tableau de contingence

On calcule les effectifs marginaux par :

$$x_{\bullet j} = \sum_{i=1}^{l} x_{ij}$$

Et la taille des échantillons dans chaque population par :

$$n_i = \sum_{j=1}^{c} x_{ij}$$

De plus, on a :

$$\sum_{i=1}^{l}\sum_{j=1}^{c} x_{ij} = \sum_{i=1}^{l} n_i = \sum_{j=1}^{c} x_{\bullet j} = n$$

2.3.2 Règle de décision

Les hypothèses du test sont :

$$\begin{cases} H_0 : \text{les } l \text{ populations sont homogènes} \\ H_1 : \text{ non } H_0 \end{cases}$$

On note p_{ij} la probabilité pour un individu de population P_i de posséder la modalité m_j pour $1 \le i \le l$ et $1 \le j \le c$.

Les populations sont homogènes si les p_{ij} ne dépendent pas de la population P_i pour tout $(i\ ;\ j)$, ce qui se traduit par : $p_{ij} = \mathrm{IP}(X = m_j)$ pour $1 \le i \le l$ et $1 \le j \le c$

Comme on ne connait pas la loi de X, les valeurs $p_j = \mathrm{IP}\left(X = m_j\right)$ sont inconnues, et on estimera p_j par $\frac{x_{\bullet j}}{n}$.

On note $n_{ij} = \frac{n_i x_{\bullet j}}{n} = n_i p_j$, appelé *effectif théorique.*

Quand H_0 est vraie, on peut estimer p_{ij} par $\frac{n_{ij}}{n_i}$.

On mesure l'écart entre les effectifs théoriques et expérimentaux par la quantité suivante :

$$\boxed{t_n = \sum_{i=1}^{l}\sum_{j=1}^{c} \frac{\left[x_{ij} - n_{ij}\right]^2}{n_{ij}}}$$

Théorème 9 Soient X_{ij} et N_{ij} les V.A. dont les réalisations sont respectivement x_{ij} et n_{ij}.
Si on suppose que l'hypothèse H_0 est vraie, la variable de décision

$$T_n = \sum_{i=1}^{c} \sum_{j=1}^{l} \frac{\left(X_{ij} - N_{ij}\right)^2}{N_{ij}}$$

suit un loi du Khi-2 à $(l - 1)$ $(c - 1)$ degrés de liberté.

Conséquence 3 On a la règle de décision suivante :

$$\boxed{\text{On rejette } H_0 \quad \text{ssi} \quad t_n > \chi^2_{(l-1)(c-1);1-\alpha}}$$

3 Tests de comparaison

3.1 Introduction

On considère deux V.A. X_1 et X_2 indépendantes, définies sur deux populations P_1 et P_2 respectivement, dépendant d'un paramètre inconnu θ_1 et θ_2 respectivement.
On veut tester l'hypothèse :

$$H_0 : \theta_1 = \theta_2 \quad \text{contre} \quad H_1 : \theta_1 \neq \theta_2$$

On dispose d'un n_1-échantillon de X_1 et d'un n_2-échantillon de X_2 qui fournissent respectivement T_1 un estimateur de θ_1 et T_2 un estimateur de θ_2 **.**

3.2 Comparaison des moyennes

On suppose que $\mathrm{IE}(X_1) = \mu_1$ et $\mathrm{IE}(X_2) = \mu_2$
On veut tester $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$.
On sait que $\overline{X_1}$ (resp. $\overline{X_2}$) est une bonne estimation de μ_1 (resp. μ_2).
On suppose que $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et que $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ ou que n_1 et $n_2 > 30$.
On suppose σ_1^2 et σ_2^2 connues.
On sait que $\overline{X_1} \sim \mathcal{N}\left(\mu_1; \frac{\sigma_1^2}{n_1}\right)$, $\overline{X_2} \sim \mathcal{N}\left(\mu_2; \frac{\sigma_2^2}{n_2}\right)$ et $\overline{X_1} - \overline{X_2} \sim \mathcal{N}\left(\mu_1 - \mu_2; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$.

Théorème 10 Si on suppose que l'hypothèse H_0 est vraie, la variable de décision :

$$T = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

suit une loi normale centrée réduite.

Conséquence 4 On a la règle de décision suivante :

$$\boxed{\text{On rejette } H_0 \text{ ssi } \frac{\left|\overline{x_1} - \overline{x_2}\right|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_{1-\alpha/2}}$$

Remarque 22 Si σ_1 et σ_2 sont inconnus (ce qui sera presque toujours le cas !), et que n_1 et n_2 sont suffisamment « grands »on les remplace dans le résultat précédent par les déviations standards s_{n_1-1} et s_{n_2-1}.
La règle de décision devient :

$$\boxed{\text{On rejette } H_0 \text{ ssi } \frac{\left|\overline{x_1} - \overline{x_2}\right|}{\sqrt{\frac{s_{n_1-1}^2}{n_1} + \frac{s_{n_2-1}^2}{n_2}}} > u_{1-\alpha/2}}$$

3.2 Comparaison des variances

On suppose que $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et que $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$.

On veut tester $H_0 : \sigma_1 = \sigma_2$ contre $H_1 : \sigma_1 \neq \sigma_2$.

On sait que $S_{n_1-1}^2$ (resp. $S_{n_2-1}^2$) est une bonne estimation de σ_1^2 (resp. σ_2^2).
On note $s_{n_1-1}^2$ et $s_{n_2-1}^2$ leurs réalisations, et on suppose que $s_{n_1-1}^2 \geq s_{n_2-1}^2$.

Théorème 11 La V.A. $\dfrac{S_{n_1-1}^2 / \sigma_1^2}{S_{n_2-1}^2 / \sigma_2^2}$ suit une loi de Fisher à (n_1 – 1 , n_2 – 1) d.d.l.
Si on suppose que l'hypothèse H_0 est vraie, la variable de décision :

$$T = \frac{S_{n_1-1}^2}{S_{n_2-1}^2}$$

suit une loi de Fisher à (n_1 – 1 , n_2 – 1) d.d.l.

Conséquence 5 On a la règle de décision suivante :

$$\boxed{\text{On rejette } H_0 \text{ ssi } \frac{s_{n_1-1}^2}{s_{n_2-1}^2} > f_{n_1-1;n_2-1;1-\alpha}}$$

où $f_{u;v;1-\alpha}$ est le fractile d'ordre $1-\alpha$ d'une loi de Fisher à (u ; v) d.d.l. c'est-à-dire $\mathbb{P}\left(F_{(u,v)} < f_{u;v;1-\alpha}\right) = 1-\alpha$ où $F_{(u,v)}$ suit une loi de Fisher à (u ; v) d.d.l

3.3 Comparaison des proportions

Soit p_1 (resp. p_2) la proportion d'individus possédant la modalité M dans la population P_1 (resp. P_2). On dispose d'un n_1-échantillon de P_1 et d'un n_2-échantillon de P_2 .

Soient F_1 et F_2 les fréquences empiriques associées à P_1 et P_2 respectivement, f_1 et f_2 leurs réalisations.

On veut tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$.

Pour cela, on utilise l'approximation d'une loi binomiale par une loi normale pour se ramener à la comparaison des moyennes.

Théorème 12 Par approximation, on a :

$$F_1 - F_2 \sim \mathcal{N}\left(p_1 - p_2 ; \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} \right).$$

Si on suppose que l'hypothèse H_0 est vraie, on a alors $p_1 = p_2 = p$, et la variable de décision :

$$T = \frac{F_1 - F_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

suit une loi normale centrée réduite.

Or, on ne connait pas p ! On le remplace donc par : $\boxed{f = \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2}}$

Conséquence 6 On a la règle de décision suivante :

$$\boxed{\text{On rejette } H_0 \text{ ssi } \frac{|f_1 - f_2|}{\sqrt{f(1-f)\left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} > u_{1-\alpha/2}}$$

.

4 En fiabilité : Test d'homogénéité d'un processus de Poisson

On a vu en cours de probabilités-fiabilité que le processus de comptage N_t suit une loi de Poisson de paramètre λt. Lorsque le paramètre λ n'est plus constant, on parle de processus de Poisson non homogène et le nouveau paramètre $\lambda(t)$ est appelé *intensité du processus*.

On observe le processus de Poisson sur un intervalle de temps $[0; t]$.
On note $T_1, \dots, T_n$ les V.A. donnant les instants de défaillances, et $t_1, \dots, t_n$ leurs réalisations.

Les hypothèses du test sont :

$$\begin{cases} H_0 : \text{le processus de Poisson est homogène} \\ H_1 : \text{ non } H_0 \end{cases}$$

Plusieurs tests sont envisageables. Nous allons en voir deux.

4.1 Test non paramétrique

On divise l'intervalle [0; t] en d intervalles de même longueur notés $I_1, \ldots, I_d$.
Soit N_k la V.A. qui donne le nombre de défaillances dans l'intervalle I_k, c'est à dire :

$$N_k = \sum_{i=1}^{n} 1_{\{T_i \in I_k\}} \quad \text{où } n = N(t)$$

On note n_k sa réalisation.

Théorème 13 Si on suppose que l'hypothèse H_0 est vraie, la variable de décision :

$$Z_n = \frac{d}{n} \sum_{k=1}^{d} \left(N_k - \frac{n}{d} \right)^2$$

suit une loi du Khi-2 à (d - 1) degrés de liberté.

Conséquence 7 On a la règle de décision suivante :

$$\boxed{\text{On rejette } H_0 \text{ ssi } \frac{d}{n} \sum_{k=1}^{d} \left(n_k - \frac{n}{d} \right)^2 > \chi^2_{d-1,1-\alpha}}$$

où $\chi^2_{d-1,1-\alpha}$ est le fractile d'ordre 1 - α de la loi du Khi-2 à (d - 1) degrés de liberté.

4.2 Test de Laplace

Si le processus de Poisson est homogène, les temps observés seront répartis de façon relativement homogène sur [0; t], alors que si λ n'est pas constant, les points seront plus concentrés vers 0 ou vers t.

Théorème 14 Si on suppose que l'hypothèse H_0 est vraie, la variable de décision :

$$L_n = \sqrt{\frac{12}{n}} \frac{\sum_{i=1}^{n} \left(T_i - \frac{t}{2} \right)}{t}$$

suit une loi normale centrée réduite.

Conséquence 8 On a la règle de décision suivante :

$$\text{On rejette } H_0 \text{ ssi } \sqrt{\frac{12}{n}} \frac{\left| \sum_{i=1}^{n} \left(t_i - \frac{t}{2} \right) \right|}{t} > u_{1-\alpha}$$

où $u_{1-\alpha}$ est le fractile d'ordre (1 - α) de la loi gaussienne centrée réduite, c'est-à-dire $\mathrm{IP}(U < u_{1-\alpha}) = 1 - \alpha$, où U suit une loi normale centrée réduite.