Modélisation Statistique
Licence 3 MIASHS – Université de Bordeaux

– Chapitre III –
Notes de cours
Initiation aux tests d'hypothèses

Nous abordons sommairement dans ce cours un aspect primordial de la statistique inférentielle, que l'on appelle le « test d'hypothèse ». Ce dernier forme un des chaînons souvent manquants entre la théorie et la pratique, il permet en effet d'évaluer, de mesurer la pertinence d'une hypothèse de départ dans une modélisation, sur la base des observations.

1 Principe

Supposons que l'on se donne un n–échantillon aléatoire sur lequel nous souhaitons formuler une ou plusieurs hypothèses de départ : gaussien, de Poisson, indépendant, d'espérance μ_0, de variance σ_0^2, etc. Ces hypothèses peuvent résulter de considérations visuelles (par exemple, l'histogramme fait apparaître une courbe en cloche, ou bien l'estimateur de l'espérance est très proche de la valeur supposée), ou encore intuitives (par exemple, l'échantillon étant de très grande taille, il semble judicieux de le considérer comme gaussien, ou bien l'expérience étant conduite sur des individus bien distincts, nous pouvons retenir l'indépendance). Sur la base d'observations, le test a pour objectif de quantifier l'évidence statistique de l'hypothèse de départ. On se pose dès lors la question suivante : si mon hypothèse de départ est vraie, alors mes observations sont-elles conformes à ce qui était censé se produire ? Il s'agit alors, pour répondre au mieux à la question, de produire une statistique dont on connaît le comportement sous l'hypothèse de départ et qui, de préférence, se comporte de manière radicalement différente lorsque l'hypothèse de départ est fausse. Nous construisons ainsi un intervalle (zone de non rejet) comprenant avec une forte probabilité (souvent, 95%) la statistique de test, si l'hypothèse de départ est valide. Lorsque l'observation sort de l'intervalle, l'écart entre la valeur attendue et la valeur observée est trop grand pour que nous donnions du crédit à l'hypothèse de départ, elle est alors rejetée. Dans le cas contraire, nous accordons le bénéfice du doute à la différence observée que nous attribuons au hasard seul et, considérant que nous n'avons pas suffisamment d'informations pour rejeter l'hypothèse de départ, nous la conservons au bénéfice du doute.

Définition 1.1 *On appelle « hypothèse statistique » un énoncé ou une affirmation concernant les caractéristiques d'une population aléatoire.*

Définition 1.2 *Un « test d'hypothèse » est une procédure statistique ayant pour but de fournir une règle de décision permettant, sur la base d'observations, de discriminer deux hypothèses statistiques.*

Exemples.

L'hypothèse de départ, sur laquelle repose notre affirmation, est aussi appelée « l'hypothèse nulle » et on la note traditionnellement $\mathcal{H}_0$. Elle est donc considérée comme vraie, jusqu'à ce qu'un test d'hypothèse permette de la rejeter en faveur d'une « hypothèse alternative » notée $\mathcal{H}_1$. En général on choisit $\mathcal{H}_1 = \bar{\mathcal{H}}_0$, mais ce n'est pas toujours le cas.

Exemples.

2 Socle théorique

Considérons un n–échantillon $(X_1, \ldots, X_n)$ sur lequel porte une hypothèse nulle $\mathcal{H}_0$ et une hypothèse alternative $\mathcal{H}_1$. Soit T_n une statistique de l'échantillon dont on connaît le comportement sous $\mathcal{H}_0$. Supposons ainsi que, pour un niveau $0 \leq \alpha \leq 1$ donné, le comportement de T_n soit caractérisé par

$$\mathbb{P}\big(q_{\alpha/2} \leq T_n \leq q_{1-\alpha/2} \mid \mathcal{H}_0\big) = 1 - \alpha$$

où $q_{\alpha/2}$ et $q_{1-\alpha/2}$ sont les quantiles d'ordre $\alpha/2$ et $1-\alpha/2$ de la loi (éventuellement limite) de T_n sous $\mathcal{H}_0$. Si l'on sait de plus montrer que T_n prend un caractère explosif sous $\mathcal{H}_1$, avec par exemple

$$|\,T_n| \xrightarrow[\mathcal{H}_1]{\mathbb{P}} +\infty \quad \text{ou bien} \quad \mathbb{E}[|\,T_n|] \xrightarrow[\mathcal{H}_1]{} +\infty,$$

ou encore que T_n est simplement décalée d'un côté ou de l'autre du spectre (c'est-à-dire l'espace dans lequel T_n prend ses valeurs, en général $\mathbb{R}$ ou $\mathbb{R}^+$), alors il est possible de construire une zone de rejet et une règle associée telles que

$$\mathbb{P}\big(\text{accepter } \mathcal{H}_0 \,|\, \mathcal{H}_0 \text{ est vraie}\big) = 1-\alpha \quad \text{et} \quad \mathbb{P}\big(\text{rejeter } \mathcal{H}_0 \,|\, \mathcal{H}_0 \text{ est vraie}\big) = \alpha.$$

Si, dans le meilleur des cas, on a montré que T_n était rejetée à l'infini sous $\mathcal{H}_1$, la zone de rejet est donnée par $\mathcal{R} =]-\infty, q_{\alpha/2}\,[\, \cup \,]\, q_{1-\alpha/2}, +\infty[$ et la zone d'acceptation par $\mathcal{A} = [q_{\alpha/2}, q_{1-\alpha/2}]$. Ainsi, une fois que la statistique T_n est observée sur le jeu de données, la règle de décision est la suivante :

— $T_n \in \mathcal{R} \implies T_n$ se trouve dans une zone où elle n'avait qu'une très faible probabilité de se trouver si l'hypothèse $\mathcal{H}_0$ était vraie. Sachant de plus que T_n a tendance à exploser sous $\mathcal{H}_1$ et donc à se trouver dans $\mathcal{R}$, nous choisissons de rejeter $\mathcal{H}_0$ en faveur de $\mathcal{H}_1$, plus plausible sur la base de nos observations.

— $T_n \in \mathcal{A} \implies T_n$ se trouve en effet dans un intervalle où, si $\mathcal{H}_0$ était vraie, elle avait de grandes chances de se trouver, ce qui ne veut pas dire pour autant que $\mathcal{H}_0$ est forcément vraie. Par manque d'informations permettant de rejeter $\mathcal{H}_0$, nous décidons de lui accorder le bénéfice du doute.

Selon cette méthodologie, T_n jouant le rôle d'arbitre dans la décision finale, on l'appelle la « statistique de test » : son comportement sous $\mathcal{H}_0$ permet d'établir l'intervalle de confiance, et son comportement sous $\mathcal{H}_1$ permet de fixer la zone de rejet.

2.1 Un exemple détaillé

Nous disposons d'une observation du n–échantillon $(X_1, \ldots, X_n)$ gaussien (c'est déjà là une hypothèse... mais pas celle que nous allons tester) de loi parente $\mathcal{N}(\mu, \sigma^2)$. Cependant, et comme son estimateur naturel est proche de μ_0, nous souhaitons effectuer le test d'hypothèse suivant : nos observations sont-elles de moyenne $\mu = \mu_0$? Nous nous plaçons donc sous l'hypothèse nulle $\mathcal{H}_0$: "$\mu = \mu_0$" et nous allons la tester contre l'hypothèse alternative $\mathcal{H}_1$: "$\mu \neq \mu_0$". En conséquence,

$$T_n = \sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \underset{\mathcal{H}_0}{\sim} \mathcal{N}(0,1).$$

Ainsi, T_n sera notre statistique de test et son comportement sous $\mathcal{H}_0$ sera décrit par la loi normale $\mathcal{N}(0,1)$. Par ailleurs, si l'on se place sous $\mathcal{H}_1$, alors $\bar{X}_n \sim \mathcal{N}(\mu_1, \sigma^2/n)$ avec $\mu_1 \neq \mu_0$. Il s'ensuit que

$$T_n \underset{\mathcal{H}_1}{\sim} \mathcal{N}\Big(\frac{\sqrt{n}\,(\mu_1 - \mu_0)}{\sigma},\, 1\Big)$$

et il est donc clair que T_n est rejetée à l'infini sous $\mathcal{H}_1$. Plus précisément, T_n est rejetée vers $-\infty$ lorsque $\mu_0 > \mu_1$ et vers $+\infty$ lorsque $\mu_0 < \mu_1$. Nous pouvons en outre tirer de la loi de T_n sous $\mathcal{H}_0$ une zone d'acceptation de niveau $1-\alpha$ dans laquelle, en supposant que l'hypothèse nulle soit valide, T_n a une grande probabilité de se trouver (une probabilité de $1-\alpha$, exactement). Ainsi,

$$\mathbb{P}\big(-\,u_{1-\alpha/2}\;\leq\;T_n\;\leq\;u_{1-\alpha/2}\mid\mathcal{H}_0\big)=1-\alpha$$

où $u_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ de la loi normale $\mathcal{N}(0,1)$. Cela nous montre que

$$\mathcal{A}=[-u_{1-\alpha/2},\,u_{1-\alpha/2}].$$

Par ailleurs, et puisque $|\,T_n|$ est rejetée à l'infini sous $\mathcal{H}_1$, on a

$$\mathcal{R}=\,]-\infty,\,-u_{1-\alpha/2}\,[\,\cup\,]\,u_{1-\alpha/2},\,+\infty\,[.$$

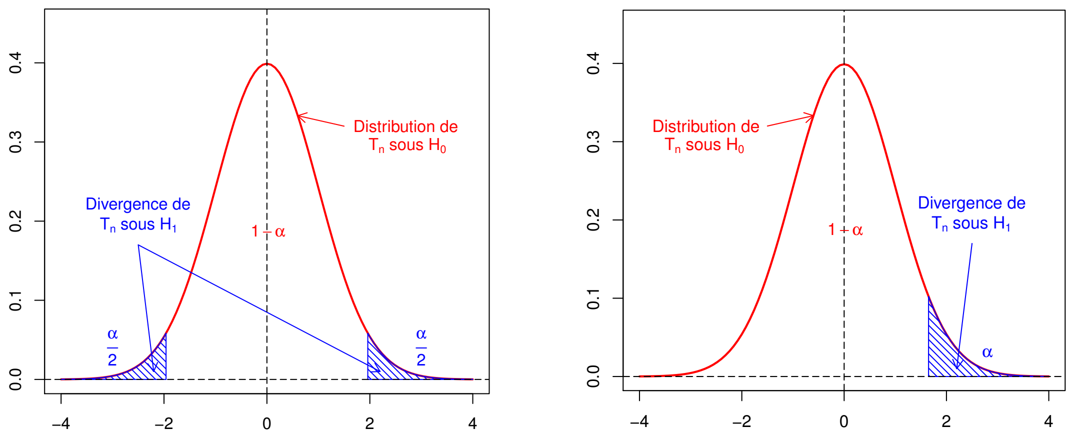
Supposons désormais que nous modifions les hypothèses à tester, pour les exprimer sous la forme $\mathcal{H}_0$: "$\mu\leq\mu_0$" contre $\mathcal{H}_1$: "$\mu>\mu_0$". Alors, le comportement de la statistique de test lorsque $\mu=\mu_0$ est toujours celui décrit précédemment. En revanche, nous voyons que seul le cas où T_n est rejetée vers $+\infty$ caractérise la zone de rejet de l'hypothèse nulle puisque le cas où $\mu<\mu_0$, pour lequel T_n est rejetée vers $-\infty$, reste acceptable sous $\mathcal{H}_0$. Nous en déduisons que

$$\mathcal{A}=\,]-\infty,\,u_{1-\alpha}]$$

et que

$$\mathcal{R}=\,]\,u_{1-\alpha},\,+\infty\,[.$$

Les graphes ci-dessous illustrent le comportement de la statistique de test sous $\mathcal{H}_0$, les zones hâchurées étant les plus favorables à l'hypothèse alternative $\mathcal{H}_1$, respectivement dans le cas du test de l'hypothèse nulle $\mathcal{H}_0$: "$\mu=\mu_0$" puis dans le cas de $\mathcal{H}_0$: "$\mu\leq\mu_0$", contre leurs alternatives décrites précédemment.



D'une manière générale, nous dirons que le test est « bilatéral » lorsque l'hypothèse nulle est simple (telle que $\mathcal{H}_0$: "$\mu=\mu_0$"), cela se traduit souvent par la présence d'une zone de rejet composite (une aire $\alpha/2$ de chaque côté du spectre). En revanche, lorsque l'hypothèse nulle est composite (telle que $\mathcal{H}_0$: "$\mu\leq\mu_0$"), alors nous dirons que le test est « unilatéral » et la zone de rejet n'englobe qu'une partie du spectre, d'aire α. Il est très facile de le visualiser sur les graphes ci-dessus.

2.2 Risques d'erreur, puissance et valeur test

2.2.1 Le niveau du test

Le « niveau de signification » du test est la valeur que l'on nomme traditionnellement α, avec $0 \leq \alpha \leq 1$. Elle fixe la probabilité de rejeter à tort l'hypothèse nulle, et donc la probabilité que la statistique de test T_n tombe dans la zone de rejet sachant que l'hypothèse nulle $\mathcal{H}_0$ était vraie. On l'appelle également « risque de première espèce » ou « risque de faux-positifs ». Elle caractérise donc la probabilité

$$\mathbb{P}\big(q_{\alpha/2} \leq T_n \leq q_{1-\alpha/2} \mid \mathcal{H}_0\big) = 1 - \alpha$$

où $q_{\alpha/2}$ et $q_{1-\alpha/2}$ sont les quantiles de la loi de T_n sous $\mathcal{H}_0$. Elle est fixée *a priori* et détermine entièrement l'ampleur de la zone d'acceptation $\mathcal{A} = [q_{\alpha/2},\, q_{1-\alpha/2}]$.

2.2.2 La puissance du test

La « puissance » du test caractérise la probabilité de rejeter $\mathcal{H}_0$ alors que $\mathcal{H}_0$ est effectivement fausse. Elle prend donc la valeur $1 - \beta$ si β est la probabilité de ne pas rejeter $\mathcal{H}_0$ sachant qu'elle est fausse. On appelle encore β le « risque de seconde espèce » ou « risque de faux-negatifs». Contrairement au niveau du test, la puissance n'est pas fixée *a priori* mais découle du comportement de la statistique de test sous l'hypothèse alternative que l'on a choisie. Les meilleurs tests sont bien évidemment les plus puissants, les plus aptes à rejeter $\mathcal{H}_0$ à raison.

2.2.3 Récapitulatif

Le tableau suivant résume la probabilité de non rejet ou de rejet de $\mathcal{H}_0$ selon la règle de décision induite par la statistique de test, et conditionnée par la connaissance de l'hypothèse de départ. Il est important de rappeler une nouvelle fois que α est fixée au préalable (en général à 0.05), tandis que β découle du choix de $\mathcal{H}_1$ et de α.

	$\mathcal{H}_0$ est vraie	$\mathcal{H}_0$ est fausse
$\mathcal{H}_0$ n'est pas rejetée	$1-\alpha$	β
$\mathcal{H}_0$ est rejetée	α	$1-\beta$

2.2.4 La p–valeur

Une règle de rejet assez facile à mettre en œuvre et couramment utilisée par les logiciels statistiques, consiste à calculer la probabilité que la statistique de test soit égale à la valeur observée ou encore plus extrême, tout en supposant que l'hypothèse nulle $\mathcal{H}_0$ est vraie : on appelle cette probabilité la « p–valeur » ou « probabilité critique ». Ainsi, comparer la p–valeur avec α est équivalent à déterminer si T_n est tombée dans la zone de rejet. Reprenons notre exemple détaillé à la Section 2.1. Puisque, sous $\mathcal{H}_0$: "$\mu = \mu_0$",

$$T_n \underset{\mathcal{H}_0}{\sim} \mathcal{N}(0,1),$$

il est clair que, pour le test bilatéral,

$$p\text{–val} = \mathbb{P}\big(|Z| \geq |T_n| \mid \mathcal{H}_0\big) = 2(1 - \Phi(|T_n|))$$

où Z suit la loi de la statistique de test sous $\mathcal{H}_0$, qui est la loi $\mathcal{N}(0,1)$ dans notre exemple, et Φ sa fonction de répartition. La statistique T_n est ici observée sur le jeu de données, il s'agit d'une réalisation de la variable aléatoire. Nous voyons donc que la p–valeur est une probabilité calculée *a posteriori*, en fonction des données. En particulier, dès que

$$p\text{–val} \leq 0.05,$$

alors il y avait moins de 5% de chance que T_n prenne la valeur observée, ce pourquoi on rejettera l'hypothèse nulle. D'une manière générale, on la rejettera lorsque p–val $\leq \alpha$, et le test est effectué au risque α. En pratique, il convient de rejeter fermement $\mathcal{H}_0$ lorsque p–val est très proche de 0 alors qu'il convient de ne pas rejeter $\mathcal{H}_0$ lorsque p–val grandit au-dessus de 0.1. Autour de 0.05, il y a toujours un certain doute et les conclusions devront être renforcées par d'autres procédures de test.

2.3 Un exemple avec R

Nous avons généré un vecteur gaussien $\mathcal{N}(2,2)$ de taille $n = 100$ avec R, en utilisant la commande :

```
> X=rnorm(100,2,2)
```

et nous avons évalué les estimateurs $\bar{X}_{100} \approx 1.7799$ et $S_n \approx 1.7632$. On souhaite alors effectuer le test de l'hypothèse nulle $\mathcal{H}_0$: "$\mu = 2$" contre son alternative $\mathcal{H}_1$: "$\mu \neq 2$", au niveau $\alpha = 0.05$. On entre :

```
> t.test(X, mu = 2, alternative = "two.sided", conf.level = 0.95)
```

et l'on obtient :

```
One Sample t-test
data: X
t = -1.2483, df = 99, p-value = 0.2149
alternative hypothesis: true mean is not equal to 2
95 percent confidence interval:
1.430031 2.129760
sample estimates:
mean of x
1.779895
```

Tentons de retrouver et d'interpréter ces résultats :

Nous lançons cette fois le test de l'hypothèse nulle $\mathcal{H}_0$: "$\mu \leq 1$" contre son alternative $\mathcal{H}_1$: "$\mu > 1$", mais avec un niveau $\alpha = 0.01$:

```
t.test(X, mu = 1, alternative = "greater", conf.level = 0.99)
```

et l'on obtient :

```
One Sample t-test
data: X
t = 4.4231, df = 99, p-value = 1.25e-05
alternative hypothesis: true mean is greater than 1
99 percent confidence interval:
1.362959 Inf
sample estimates:
mean of x
1.779895
```

Tentons également de retrouver et d'interpréter ces résultats :

3 Les tests d'hypothèse usuels

En vertu de ce que nous venons de développer, il est possible de dériver autant de tests d'hypothèses que l'on souhaite, dès que l'on sait identifier deux hypothèses antagonistes et le comportement d'une statistique sous l'hypothèse nulle. Nous en donnons quelques exemples classiques, très loins d'être exhaustifs, sur les échantillons indépendants et identiquement distribués. Dans toute la suite, nous notons u_α, t_α, z_α et f_α respectivement les quantiles d'ordre α des lois normale, de Student, du khi-deux et de Fisher.

3.1 Test de la moyenne

Que l'échantillon soit gaussien ou de grande taille, la statistique de test à utiliser pour tester $\mathcal{H}_0$: "$\mu = \mu_0$" contre $\mathcal{H}_1$: "$\mu \neq \mu_0$" est

$$T_n = \sqrt{n}\, \frac{\bar{X}_n - \mu_0}{\sigma}.$$

Ainsi, T_n suit la loi normale $\mathcal{N}(0,1)$ lorsque l'échantillon est gaussien, alors qu'elle converge seulement en loi vers la loi normale $\mathcal{N}(0,1)$, par l'intermédiaire du TLC, en l'absence de gaussianité. Dans un cas comme dans l'autre, $|T_n|$ est rejetée à l'infini sous $\mathcal{H}_1$ comme nous l'avons vu précédemment, et la zone de rejet est donnée par

$$\mathcal{R} =]-\infty, -u_{1-\alpha/2}[\,\cup\,]u_{1-\alpha/2}, +\infty[.$$

Lorsque le test est unilatéral et que l'on teste $\mathcal{H}_0$: "$\mu \leq \mu_0$" contre son alternative usuelle, la zone de rejet est

$$\mathcal{R} =]u_{1-\alpha}, +\infty[.$$

Si maintenant la variance σ^2 est inconnue, on la remplace par son estimateur débiaisé S_n^2. La statistique de test devient alors

$$T_n = \sqrt{n}\, \frac{\bar{X}_n - \mu_0}{S_n}$$

et la zone de rejet reste la même pour les grands échantillons. Dans le cas gaussien, il convient de remplacer les quantiles par ceux de la loi $t(n-1)$, comme nous l'avons vu dans le Chapitre II.

3.2 Test de la variance

Pour les échantillons gaussiens de loi $\mathcal{N}(\mu, \sigma^2)$, on a par exemple le test d'intérêt pratique $\mathcal{H}_0$: "$\sigma^2 \leq \sigma_0^2$" contre $\mathcal{H}_1$: "$\sigma^2 > \sigma_0^2$", qui se traduit par la recherche d'une valeur limite pour la variance d'une population d'espérance inconnue. La statistique de test est alors

$$T_n = \frac{n-1}{\sigma_0^2}\, S_n^2.$$

La statistique T_n suit donc la loi $\chi^2(n-1)$ sous l'hypothèse nulle, alors que l'espérance de la loi limite est amplifiée par $\sigma^2/\sigma_0^2 > 1$ sous l'alternative, ce qui tend à rejeter T_n à la droite du spectre, et ce d'autant plus vite que σ^2 est grand. On isole ainsi

$$\mathcal{R} =]z_{n-1,\,1-\alpha}, +\infty[.$$

Si au contraire on considère le test bilatéral avec $\mathcal{H}_0$: "$\sigma^2 = \sigma_0^2$", alors il faut également tenir compte du rapport $\sigma^2/\sigma_0^2 < 1$ qui tend à rejeter T_n à la gauche du spectre, la zone de rejet s'écrit alors

$$\mathcal{R} = [0, z_{n-1,\,\alpha/2}[\,\cup\,]z_{n-1,\,1-\alpha/2}, +\infty[.$$

Enfin, dans le cas plus rare où l'on considère que l'espérance μ est connue, on ne doit plus estimer σ^2 par s_n^2 mais par

$$\tilde{s}_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)^2.$$

Il est alors facile de montrer que la statistique de test

$$T_n = \frac{n}{\sigma_0^2}\,\tilde{s}_n^2$$

engendre la même zone de rejet avec les quantiles de la loi $\chi^2(n)$.

3.3 Test de la proportion

Lorsque le n–échantillon est de loi parente $\mathcal{B}(p)$, il est également possible de construire le test $\mathcal{H}_0$: "$p = p_0$" contre $\mathcal{H}_1$: "$p \neq p_0$", à partir de la statistique de test

$$T_n = \sqrt{n}\frac{\bar{X}_n - p_0}{\sqrt{p_0\,(1-p_0)}}.$$

On montre en effet par le TLC que, sous $\mathcal{H}_0$,

$$T_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

alors que, sous $\mathcal{H}_1$,

$$|\,T_n\,| \xrightarrow{\mathbb{P}} +\infty.$$

On obtient ainsi immédiatement

$$\mathcal{R} =]-\infty, -u_{1-\alpha/2}[\ \cup\ ]u_{1-\alpha/2}, +\infty[$$

et nous pouvons en déduire également la zone de rejet associée aux tests unilatéraux tels que $\mathcal{H}_0$: "$p \leq p_0$", à partir du quantile $u_{1-\alpha}$.

3.4 Comparaison de deux échantillons

3.4.1 Égalité des moyennes

Soient les échantillons $(X_1, \dots, X_m)$ et $(Y_1, \dots, Y_n)$, mutuellement indépendants et de lois parentes respectives $\mathcal{N}(\mu_1, \sigma^2)$ et $\mathcal{N}(\mu_2, \sigma^2)$. On considère l'estimateur de la variance de l'échantillon global,

$$S_n^2 = \frac{m-1}{m+n-2}\,S_X^2 + \frac{n-1}{m+n-2}\,S_Y^2$$

où S_X^2 et S_Y^2 sont les estimateurs débiaisés de σ^2 sur chacun des échantillons. Soit la statistique de test

$$T_{n,m} = \sqrt{mn}\,\frac{\bar{X}_m - \bar{Y}_n}{S_n\sqrt{m+n}}.$$

Alors, il est possible de montrer que, sous $\mathcal{H}_0$: "$\mu_1 = \mu_2$",

$$T_{n,m} \sim t(m+n-2)$$

alors que, sous $\mathcal{H}_1$: "$\mu_1 \neq \mu_2$", $|T_{n,m}|$ est rejetée à l'infini lorsque n et m grandissent, le numérateur n'étant plus centré. La zone de rejet est donnée par

$$\mathcal{R} =]-\infty, -t_{m+n-2,1-\alpha/2}[\ \cup\ ]t_{m+n-2,1-\alpha/2}, +\infty[.$$

Dans le cas où σ^2 est connue, il suffit de remplacer S_n par σ dans la statistique de test, et la loi sous $\mathcal{H}_0$ reste la loi normale $\mathcal{N}(0,1)$. La zone de rejet en est immédiatement déduite. Lorsque l'alternative est $\mathcal{H}_1$: "$\mu_1 < \mu_2$", il est facile de voir que la statistique T_n est rejetée vers la gauche du spectre : on utilise alors le quantile d'ordre $t_{m+n-2,1-\alpha}$, et la zone de rejet devient

$$\mathcal{R} =]-\infty, -t_{m+n-2,1-\alpha}[.$$

3.4.2 Égalité des variances

Soient les échantillons $(X_1, \ldots, X_m)$ et $(Y_1, \ldots, Y_n)$, mutuellement indépendants et de lois parentes respectives $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$. On considère la statistique de test

$$T_n = \frac{S_X^2}{S_Y^2}$$

où S_X^2 et S_Y^2 sont les estimateurs débiaisés de σ_1^2 et de σ_2^2, sur chacun des échantillons. On a vu dans le Chapitre II que

$$\frac{\sigma_2^2}{\sigma_1^2}\, T_n \sim F(m-1, n-1).$$

Il s'ensuit que, sous $\mathcal{H}_0$: "$\sigma_1^2 = \sigma_2^2$",

$$T_n \sim F(m-1, n-1)$$

alors que, sous $\mathcal{H}_1$: "$\sigma_1^2 \neq \sigma_2^2$", l'espérance de T_n est décalée proportionnellement à l'ampleur de σ_1^2/σ_2^2. La zone de rejet est donnée par

$$\mathcal{R} = [0, f_{m-1,n-1,\alpha/2}[\ \cup\ ]f_{m-1,n-1,1-\alpha/2}, +\infty[.$$

Là encore, selon la valeur du rapport σ_1^2/σ_2^2, il est possible de discriminer une hypothèse unilatérale telle que $\mathcal{H}_1$: "$\sigma_1^2 > \sigma_2^2$", à l'aide du quantile $f_{m-1,n-1,1-\alpha}$.

3.5 Test de l'indépendance

On considère ici deux variables aléatoires X et Y qualitatives, et l'on souhaite remettre en question le fait que ces deux variables sont indépendantes, dans le sens où une réalisation de Y n'est pas influencée par la connaissance de la variable X correspondante, par l'intermédiaire d'un test d'hypothèse. On construit ainsi un tableau de contingence avec en lignes les modalités de X (que l'on appelle A_1, ..., A_j), en colonnes celles de Y (que l'on appelle B_1, ..., B_k), et dans chaque case les effectifs $(n_{h,\ell})$, pour tout $1 \leq h \leq j$ et $1 \leq \ell \leq k$, correspondant à un n–échantillon observé.

	B_1	B_2	$\dots$	B_{k-1}	B_k	Somme
A_1	$n_{1,1}$	$n_{1,2}$	$\dots$	$n_{1,k-1}$	$n_{1,k}$	n_{A_1}
A_2	$n_{2,1}$	$n_{2,2}$	$\dots$	$n_{2,k-1}$	$n_{2,k}$	n_{A_2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_{j-1}	$n_{j-1,1}$	$n_{j-1,2}$	$\dots$	$n_{j-1,k-1}$	$n_{j-1,k}$	$n_{A_{j-1}}$
A_j	$n_{j,1}$	$n_{j,2}$	$\dots$	$n_{j,k-1}$	$n_{j,k}$	n_{A_j}
Somme	n_{B_1}	n_{B_2}	$\dots$	$n_{B_{k-1}}$	n_{B_k}	n

Sous l'hypothèse nulle $\mathcal{H}_0$: "X et Y sont indépendantes", on s'attend à ce que l'effectif de chaque croisement soit engendré par proportionnalité entre les effectifs totaux. On construit ainsi les effectifs théoriques $(p_{h,\ell})$ tels que, pour tout $1 \leq h \leq j$ et $1 \leq \ell \leq k$,

$$p_{h,\ell} = \frac{n_{A_h} n_{B_\ell}}{n}.$$

On calcule alors la statistique de test

$$T_n = \sum_{h=1}^{j} \sum_{\ell=1}^{k} \frac{(n_{h,\ell} - p_{h,\ell})^2}{p_{h,\ell}}$$

et l'on peut montrer que sous $\mathcal{H}_0$, T_n est bien approximée par la loi $\chi^2((j-1)(k-1))$ pour n assez grand, alors que T_n est rejetée à l'infini sous l'hypothèse alternative $\mathcal{H}_1$ stipulant que X et Y ne sont pas indépendantes. On en déduit la zone de rejet

$$\mathcal{R} =]z_{(j-1)(k-1),\,1-\alpha}, +\infty[.$$

Notons pour conclure que, dans le cas où les variables X et Y sont continues, il est toujours possible, si cela semble pertinent dans le cadre de l'étude, de les discrétiser en classes/effectifs et de les traiter comme des variables qualitatives.

4 Résumé

Mener une procédure de test d'hypothèse, c'est :

- Identifier les variables aléatoires et les hypothèses de départ.
- Identifier une hypothèse $\mathcal{H}_0$ à remettre en question et son alternative $\mathcal{H}_1$.
- Identifier une statistique de test T_n facile à manipuler sous $\mathcal{H}_0$.
- Identifier le comportement de T_n sous $\mathcal{H}_0$:
  - À l'aide de la loi parente : comportement exact.
  - À l'aide du TLC : comportement asymptotique.
- Déduire le comportement de T_n sous $\mathcal{H}_1$:
  - De préférence rejetée vers l'infini.
- Fixer un risque de première espèce α :
  - Souvent à 0.05, voire à 0.01.
- Isoler une zone de rejet $\mathcal{R}$ de probabilité α sous $\mathcal{H}_0$:
  - En isolant la région du spectre où T_n est censée se trouver sous $\mathcal{H}_1$.
- Évaluer la puissance $1 - \beta$ du test, son meilleur critère de qualité :

- Parfois impossible, si l'on ne dispose pas de la distribution de T_n sous $\mathcal{H}_1$.

— Observer un échantillon et calculer T_n.

— Conclure :
  - $T_n \in \mathcal{R}$: $\mathcal{H}_0$ rejetée.
  - $T_n \notin \mathcal{R}$: $\mathcal{H}_0$ conservée au bénéfice du doute.