Chapitre 7

Tests d'hypothèses
Données qualitatives

7.1 Introduction

Au chapitre 6, les observations d'un échantillon X_1, X_2, …, X_n représentent des quantités mesurées — des poids, des longueurs, des résultats de tests psychométriques, etc. C'étaient donc des données *quantitatives*. Nous passons maintenant aux données dites *qualitatives* ou *catégorielles*: chaque élément de l'échantillon est *catégorisé*, et les observations seront des décomptes: le nombre d'observations appartenant à telle ou telle catégorie. Les observations seront de loi hypergéométrique, binomiale ou multinomiale; et les paramètres seront des *proportions*.

Les deux premiers tests sont une version qualitative de tests sur une moyenne: on teste l'hypothèse qu'une proportion (ou probabilité) p est égale à une valeur donnée p_o; et puis que la différence $\delta = p_1 - p_2$ entre deux proportions (ou probabilités) est égale à une valeur donnée δ_o. Ces mêmes tests sont ensuite généralisés à des tests concernant une série de proportions (*test d'ajustement*) et une série de différences de proportions (*test d'indépendance*). Finalement, on reviendra à la différence entre deux proportions dans le cas de données appariées.

L'exemple suivant illustre une approche informelle à un test concernant une proportion:

Exemple 7.1.1 *Test d'hypothèse sur un effectif—population finie*

Une compagnie ayant conclu une entente d'achat d'un bloc de 80 appartements se propose d'examiner l'état des appartements à partir d'un échantillon de 20 unités tirées au hasard. Le vendeur affirme que 52 des unités sont dans un état impeccable (ne nécessitent pas de réparations majeures). L'acheteur a le droit d'annuler le contrat s'il est établi (à un certain degré de confiance que le nombre d'unités parfaites est inférieur à 52. L'échantillon est tiré, les appartements examinés et le nombre d'unités en état impeccable est compté: 3 seulement. Peut-on affirmer avec confiance que le nombre d'unités impeccables dans la population est inférieur à 52?

Solution La population est de taille $N = 80$ et l'échantillon de taille $n = 12$. Le nombre X d'unités impeccables dans l'échantillon est une variable de loi hypergéométrique H(n; N_1, N_2), où N_1 est le nombre d'appartements impeccables et $N_2 = N - N_1$. Selon l'affirmation du vendeur, $N_1 = 52$, ce qui voudrais dire que la proportion p d'appartements impeccables dans la population est $p = 0{,}65$. Dans l'échantillon, l'observation $X = 12$ correspond à une proportion observée de 3/12 = 0,25. Utilisant la fonction de probabilité de la loi hypergéométrique, on évalue la probabilité d'une valeur de X aussi petite que 3: $P(X \leq 3 \mid N_1 = 52) = 0{,}0028$, une valeur extrêmement petite qui permet d'affirmer avec confiance que $N_1 < 52$. ■

L'exemple ci-dessus est un test d'hypothèse sur le paramètre N_1, mais il est équivalent à un test sur une proportion, soit $p = \frac{N_1}{N}$. L'hypothèse nulle est

$$H_o : N_1 = 52 \text{ ou } p = \frac{52}{80} \text{ et l'alternative est } H_1\text{: } N_1 < 52 \text{ ou } p < \frac{52}{80} .$$

Ce qui est calculé est la *valeur p*, $vp = 0{,}0028$. On peut effectuer le test de façon plus formelle, une approche qui pourrait être stipulée dans l'entente.

Calcul par Excel: $P(X \leq \mathtt{x} \mid n = \mathtt{n};\ N_1 = \mathtt{N1};\ N = \mathtt{N})$ `=LOI.HYPERGEOMETRIQUE.N(x;n;N1;N;1).`

Exemple 7.1.2 (*suite de l'exemple* 7.1.1)

Par exemple, il pourrait être entendu, avant le tirage de l'échantillon, que le contrat peut être annulé seulement s'il est établi à un niveau de 5 % que $N_1 < 52$. Déterminer une région critique de niveau $\alpha = 0{,}05$ en termes de X

Solution Une région critique de la forme $X \le C$ doit satisfaire la condition $P(X \le C \mid N_2 = 52) \le 0{,}05$. On choisira la plus grande région critique qui satisfait cette condition. Étant donné que X est une variable discrète, on peut procéder par tâtonnement: on détermine la taille de plusieurs régions critiques. Les voici pour C = 0 à 5.

C	0	1	2	3	4	5
$P(X \le C \mid N_2 = 52)$	0,000000505	0,000019	0,000308	0,002841	0,0168	0,0679

La région critique est donc $X \le 4$. Sa taille est 0,0168, nettement inférieure au seuil $\alpha = 0{,}05$. Ceci est dû au fait que X est une variable discrète: la région critique $X \le 5$ est de taille supérieure au seuil. Plus n est grand, cependant, plus on peut s'approcher du niveau α ■

Remarque *Une «proportion» p n'est pas toujours un quotient concret comme dans les exemples précédents. Il représente plutôt une probabilité. Par exemple, la proportion p de pièces défectueuses produites par un procédé manufacturier est en fait la probabilité qu'une pièce soit défectueuse, une mesure de la qualité du processus de fabrication. C'est le cas de l'exemple suivant, où la population est essentiellement infinie, ce qui justifie l'emploi de la loi binomiale plutôt que la loi hypergéométrique.* ■

Calcul par Excel de la probabilité $P(X \le \texttt{x} \mid n = \texttt{n};\ p = \texttt{p};\ N = \texttt{N})$: `=LOI.BINOMIALE.N(x;n;p;1`.

Exemple 7.1.3 *Test sur une proportion – Population infinie ou très grande*

Un procédé de fabrication de boulons est considéré satisfaisant si le taux de défectuosité est de 2 % (ou moins, bien sûr). Un inspecteur prélève un échantillon de 200 boulons pour s'assurer que le procédé est bien sous contrôle. Soit X le nombre de boulons défectueux qu'on trouvera dans l'échantillon. Déterminer une région critique de niveau $\alpha = 0{,}05$ pour tester l'hypothèse

$$H_o : p = 0{,}02$$

Solution $X \sim \mathcal{B}(200\ ;\ p)$. Si H_o est vraie, $E(X) = 200(0{,}02) = 4$. On rejettera H_o si $X \ge C$, où C est le plus petit nombre supérieur à 4 pour lequel

$$P(X \ge C \mid p = 0{,}02) \le 0{,}05.$$

Le tableau suivant présente la probabilité $P(X \ge C \mid p = 0{,}02)$ pour quelques valeurs de C.

C	5	6	7	8	9	10	11
$P(X \ge C \mid p = 0{,}02)$	0,3712	0,2133	0,1086	0,0493	0,0202	0,0075	0,0025

La région critique de niveau 0,05 est donc $\{X \ge 8\}$. Elle est de taille 0,0493. ■

Dans les derniers exemples, les tests ont été exécutés à partir de la distribution exacte de l'observation X, soit la loi hypergéométrique (si la population est finie) soit la loi binomial (si la population est infinie, ou très grande). Si la taille de la population N est très grande, la loi hypergéométrique peut être approchée par la loi binomiale. Et si la taille de l'échantillon est grande (sans trop approcher N), la loi binomiale peut être approchée par la loi normale. Dans la plupart des applications réelles, ces conditions sont réunies: la population est très grande par rapport à l'échantillon (sinon infinie) et l'échantillon est assez grand pour que la loi normale soit utilisée pour approcher la loi binomiale.

7.2 Tests sur une proportion - Grands échantillons

L'exemple 7.1.3 traite d'un test d'hypothèse concernant une proportion. Lorsque n est grand, nous pouvons nous servir de la loi normale comme approximation de la distribution de $\hat{p} = X/n$. La loi de X, une loi $\mathcal{B}(n\ ;\ p)$, peut être approchée par une loi $\mathcal{N}(np\ ;\ np(1\text{-}p))$, et donc

$$\hat{p} \sim N\left(p\ ;\ \frac{p(1-p)}{n} \right) \qquad (7.2.1)$$

Considérons l'hypothèse

$$H_o : p = p_o \qquad (7.2.2)$$

et l'alternative

$$H_1 : p \neq p_o. \qquad (7.2.3)$$

La région critique peut s'exprimer en fonction de la statistique centrée-réduite

$$Z = \frac{\hat{p} - p_o}{\sigma_{\hat{p}}} = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}}. \qquad (7.2.4)$$

Lorsque H_o est vraie, Z suit à peu près une loi $N(0\ ;\ 1)$. On rejette donc H_o si

$$\frac{|\hat{p} - p_o|}{\sqrt{\frac{p_o(1-p_o)}{n}}} > z_{\alpha/2}. \qquad (7.2.5)$$

Remarque *Le dénominateur de Z est l'écart-type de* $\sigma_{\hat{p}}$ *de* $\hat{p}$ *lorsque* H_o *est vraie. Mais on pourrait remplacer le dénominateur par une estimation de l'écart-type et baser le test sur la statistique*

$$\hat{Z} = \frac{\hat{p} - p_o}{\hat{\sigma}_{\hat{p}}} = \frac{\hat{p} - p_o}{\sqrt{\hat{p}(1-\hat{p})/n}} \qquad (7.2.6)$$

qui, elle aussi, suit à peu près une loi $N(0\ ;\ 1)$ *lorsque n est grand et donc la région critique est* $|\hat{Z}| > z_{\alpha/2}$ *:*

On rejette H_o si

$$\left| \frac{\hat{p} - p_o}{\sqrt{\hat{p}(1-\hat{p})/n}} \right| > z_{\alpha/2}. \qquad (7.2.6')$$

Les deux tests ne sont pas équivalents et peuvent théoriquement mener à des conclusions différentes. Mais cette éventualité est peu probable lorsque n est grand. ■

Pour des tests unilatéraux, on a les régions critiques suivantes :

$Z > z_\alpha$ si l'alternative est $H_1' : p > p_o$

$Z < -z_\alpha$ si l'alternative est $H_1'' : p < p_o$

Exemple 7.2.1 *Test d'hypothèse sur une proportion – calculs par approximation normale*

Un grossiste conclut avec son fournisseur de noix une entente selon laquelle 30 % des noix de chaque lot livré doivent être des noix de cajou. Le grossiste prévoit prélever un échantillon de 100 noix à chaque livraison afin de s'assurer que la proportion de noix de cajou ne soit pas inférieure à 30 %. Lors d'une livraison, on trouve seulement 20 noix de cajou dans un échantillon de 100 noix. Déterminer un test unilatéral à 5 % et effectuer le test.

Solution Si p est la proportion de noix de cajou dans le lot, l'hypothèse nulle est $H_o : p = 0,3$, l'alternative $H_1 : p < 0,3$. La région critique est $Z < -1,645$, où $Z = \dfrac{\hat{p}-p_o}{\sqrt{\dfrac{p_o(1-p_o)}{n}}}$.

$Z = \dfrac{0,2-0,3}{\sqrt{\dfrac{0,3(1-0,3)}{100}}} = -2,18 < -1,645$. Donc on rejette H_o. ■

Exemple 7.2.2 *Test sur une proportion — qualité de l'approximation normale*

Dans l'exemple 7.2.1, comparer $\hat{Z} = \dfrac{\hat{p}-p_o}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}}$ à $Z = \dfrac{\hat{p}-p_o}{\sqrt{\dfrac{p_o(1-p_o)}{n}}}$ et les valeurs p correspondantes (telles qu'approchées par la loi normale). Examiner la qualité des approximations normales en comparant les valeurs p d'un test basé sur $\hat{Z}$, d'un test basé sur Z et d'un test exact basé sur la loi binomiale à la valeur p d'un test basé sur X = nombre de noix de cajou.

Solution $\hat{Z} = -2,5$, et $vp = P(\hat{Z} \le -2,5 \mid H_o) = 0,0062$; $Z = -2,18$ et $vp = P(Z \le -2,18 \mid H_o) = 0,0145$. X est de loi $\mathcal{B}(100; 0,3)$ sous H_o et la valeur p *exacte* est $P(X \le 20 \mid H_o) = 0,0165$. On constate que le test basé sur Z est ici plus précis que celui basé sur $\hat{Z}$. ■

7.3 Tests sur la différence entre deux proportions

Considérons deux variables binomiales indépendantes, X_1 et X_2, où $X_1 \sim \mathcal{B}(n_1 ; p_1)$ et $X_2 \sim \mathcal{B}(n_2 ; p_2)$. La statistique pour tester l'hypothèse

$$H_o : p_1 = p_2$$

contre

$$H_1 : p_1 \ne p_2$$

serait, idéalement,

$$\frac{\hat{p}_1-\hat{p}_2}{\sigma_{\hat{p}_1-\hat{p}_2}} \qquad (7.3.1)$$

où

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}} \qquad (7.3.2)$$

On estimerait donc $\sigma_{\hat{p}_1-\hat{p}_2}$ par

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \qquad (7.3.3)$$

Cependant, si on tient compte de H_o ($p_1 = p_2 = p$ disons), alors

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p(1-p)}{n_1}+\frac{p(1-p)}{n_2}} = \sqrt{p(1-p)}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \qquad (7.3.4)$$

et alors on estimera $\sigma_{\hat{p}_1-\hat{p}_2}$ par

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \qquad (7.3.5)$$

où $\hat{p}$ est une estimation de la proportion commune aux deux populations.

Si $p_1 = p_2$, alors $X + Y$ est une variable de loi $\mathcal{B}(n_1+n_2\,;\,p)$ et p est alors estimée par

$$\hat{p} = \frac{X+Y}{n_1+n_2} = \frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2} \qquad (7.3.6)$$

La statistique

$$Z = \frac{\hat{p}_1-\hat{p}_2}{\hat{\sigma}_{\hat{p}_1-\hat{p}_2}}$$

suit à peu près une loi $\mathcal{N}(0\,;\,1)$ sous H$_o$, quelle que soit la façon dont $\sigma_{\hat{p}_1-\hat{p}_2}$ est estimée. On rejette H$_o$ si $|Z| > z_{\alpha/2}$. Pour des tests unilatéraux, on a les régions critiques suivantes :

$Z > z_\alpha$ si l'alternative est $\mathrm{H}_1' : p_1 > p_2$

$Z < -z_\alpha$ si l'alternative est $\mathrm{H}_1'' : p_1 < p_2$.

Dans la suite de ce manuel, le dénominateur de Z sera (7.3.5).

Exemple 7.3.1 *Égalité de deux proportions*

À l'occasion d'un sondage d'opinion mené auprès d'étudiants de l'UQÀM, on s'est intéressé à savoir si l'opposition à un mariage interconfessionnel émane d'une opposition généralisée au mariage à l'extérieur du groupe. Ce qui aurait pour implication que ceux qui s'opposent au mariage interconfessionnel tendent à s'opposer également au mariage interracial. Parmi les 90 répondants, 33 se sont dits opposés au mariage inter-religieux, contre 57 qui ne s'y opposaient pas. Dans le premier groupe, 22 s'opposaient aussi au mariage interracial, contre 10 dans le deuxième groupe. Peut-on conclure que ceux qui s'opposent au mariage inter-religieux sont plus nombreux à s'opposer aussi au mariage interracial ?

Solution On a deux échantillons, de tailles $n_1 = 33$ et $n_2 = 57$. Soit X_1 et X_2 le nombre de personnes qui s'opposent au mariage interracial dans le premier et deuxième groupe, respectivement. $X_1 \sim \mathcal{B}(33\,;\,p_1)$, $X_2 \sim \mathcal{B}(57\,;\,p_2)$. L'hypothèse nulle H$_o$ (que les deux groupes ne diffèrent pas quand à leur attitude sur le mariage interracial) est H$_o$: $p_1 = p_2$. On prendra pour alternative l'hypothèse H$_1$: $p_1 \neq p_2$, même si on soupçonne fortement que p_1 ne peut pas être inférieur à p_2 : on le soupçon-ne, mais on s'abstient de l'affirmer d'avance. On calcule : $\hat{p}_1 = 22/33 = 0{,}6667$ et $\hat{p}_2 = 10/57 = 0{,}1754$; $\hat{p} = \frac{22+10}{33+57} = 0{,}3556$; $\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 0{,}1047$; $Z = 4{,}69$. Il est certain, donc, que l'attitude envers le mariage interracial est liée à l'attitude envers le mariage interconfessionnel. ■

Remarque *Si au dernier exemple on avait estimé* $\sigma_{\hat{p}_1-\hat{p}_2}$ *par la formule avec*

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{(22/33)(11/33)}{33}+\frac{(10/57)(47/57)}{57}} = 0{,}0963,$$

on aurait eu $Z = 5{,}10$, *ce qui n'aurait pas changé la conclusion.* ■

7.4 Tests d'ajustement

Le test de l'hypothèse $H_o : p = p_o$ est basé sur une observation X de loi binomiale de paramètres n et p. Ce test est utile lorsque la variable observée est dichotomique: une variable qualitative qui n'a que deux modalités. Mais une variable qualitative a souvent plus de deux modalités, donc plusieurs proportions à tester—et qu'on souhaite tester simultanément. On effectue alors un test dit d'*ajustement.*

Un test d'ajustement est un test basé sur un vecteur $\boldsymbol{X} = [X_1; \ldots; X_k]$ de loi multinomiale de paramètres n et $[p_1; \ldots ; p_k]$. L'hypothèse à tester porte simultanément sur toutes les composantes du vecteur de probabilités. Le cas particulier $k = 2$ est familier : c'est le test sur p discuté à la section 7.2.

Exemple 7.4.1 *Le test sur une proportion est un test d'ajustement*

Si on observe 100 naissances pour tester l'hypothèse que la probabilité d'avoir une fille est 0,5, les données peuvent être présentées sous la forme d'une distribution comme celle-ci :

	Filles	Garçons	Total
Effectifs	$X_1 = 48$	$X_2 = 52$	100

Bien qu'il y ait deux variables, X_1 et X_2, dans le décor, l'une d'elles est redondante : si X_1 est connu, X_2 est totalement déterminé. X_1 suit une loi binomiale de paramètres $n = 100$ et p et l'hypothèse est $H_o : p = 1/2$. La statistique de test, $Z = \dfrac{X_1/100 - 1/2}{\sqrt{\dfrac{(1/2)(1-1/2)}{100}}} = \dfrac{X_1 - 50}{\sqrt{100(1/2)(1-1/2)}}$, mesure l'écart entre X_1 et 50, son espérance sous H_o. On rejette H_o si cet écart, en valeur absolue, est trop grand, donc si $|Z| > z_{\alpha/2}$ ou, ce qui est équivalent, si $Z^2 > z^2_{\alpha/2}$.

Le test est basé sur le fait que la statistique Z est à peu près de loi $\mathcal{N}(0 ; 1)$, ce qui entraîne que Z^2 est de loi χ^2_1 (à peu près). Donc la règle qui consiste à rejeter H_o si $Z^2 > \chi^2_{1;\alpha}$ est un test de niveau α. Les deux critères de rejet, $|Z| > z_{\alpha/2}$ et $Z^2 > \chi^2_{1;\alpha}$ sont équivalents car $z^2_{\alpha/2} = \chi^2_{1;\alpha}$. ■

Nous commençons par décrire la procédure, sans justification, à l'aide d'un exemple, après quoi nous développerons la théorie et le modèle sous-jacents.

Hypothèse nulle, distributions observée et théorique — un exemple

Considérons les données suivantes sur un échantillon de 780 suicides prélevé afin de déterminer si le jour de la semaine a une influence sur les suicides. L'hypothèse à tester est

H_o : La fréquence des suicides est la même pour tous les jours de la semaine.

On classe les suicides en 4 catégories, selon le moment de la semaine où le suicide a eu lieu : le début de la semaine (lundi), le milieu de la semaine (mardi à jeudi), la fin de la semaine (vendredi), et le week-end (samedi-dimanche). :

	Jour				
	Lundi	Mardi-jeudi	Vendredi	Samedi-dimanche	Total
Effectif observé	110	320	100	250	780

En termes quantitatifs, H_o s'exprime comme ceci :

H_o : La fréquence des suicides est la même pour tous les jours de la semaine

La distribution ci-dessus, appelée *distribution observée*, doit être comparée à la distribution *attendue* sous H_o. Il est clair que si H_o est vraie, les 780 suicides devraient se répartir de façon proportionnelle au nombre de jours : 1/7 le lundi ; 3/7 les mardi-jeudi ; 1/7 le vendredi ; et 2/7 les samedi-dimanche. Le tableau suivant présente cette distribution, appelée *distribution théorique* :

	Jour				
	Lundi	Mardi-jeudi	Vendredi	Samedi-dimanche	Total
Effectif théorique	$\frac{1}{7}780 = 111{,}4$	$\frac{3}{7}780 = 334{,}3$	$\frac{1}{7}780 = 111{,}4$	$\frac{2}{7}780 = 222{,}9$	780

Nous comparerons donc la distribution observée à la distribution théorique. Si l'écart est important, on rejettera H_o.

Remarque *Le même raisonnement est à la base de tout test statistique. Une valeur postulée du paramètre donne lieu à une valeur attendue de l'observation. Dans l'exemple* 7.1.1, *la valeur postulée $p = 1/2$ entraîne une valeur attendu de X_1, soit* 50. *La statistique Z mesure l'écart entre la valeur observée X_1 et la valeur attendue. L'hypothèse $p = 1/2$ est rejetée si l'écart est trop important. C'est encore ce que nous ferons dans ce chapitre, sauf qu'il faut maintenant mesurer l'écart entre un* vecteur *d'observations et un* vecteur *de valeurs attendues.* ■

Nous commençons donc par définir une mesure de l'écart entre une série d'effectifs et une autre.

Écart entre les effectifs observés et théoriques

L'écart entre les deux tableaux est mesuré par une quantité χ^2 et définie par

$$\chi^2 = \sum_i \frac{(O_i - T_i)^2}{T_i}, \qquad (7.4.1)$$

où les O_i sont les effectifs observés et les T_i sont les effectifs théoriques.

Dans l'exemple, on a

$$\chi^2 = \frac{(110-111{,}4)^2}{111{,}4} + \frac{(320-334{,}3)^2}{334{,}3} + \frac{(100-111{,}4)^2}{111{,}4} + \frac{(250-222{,}9)^2}{222{,}9} = 5{,}09.$$

Région critique

La région critique est exprimée en fonction de χ^2. On rejettera H_o si χ^2 est *trop grand* car un grand χ^2 reflète une grande différence entre les deux tableaux. La région critique sera donc de la forme

$$\chi^2 \geq C.$$

La constante C doit être choisie de telle sorte que la probabilité d'une erreur de première espèce soit égale (à peu près) à un certain nombre α. Il faut donc que

$$P(\chi^2 \geq C \mid H_o) = \alpha.$$

Pour déterminer la valeur de C, il faut connaître la distribution de χ^2 sous H_o.

Théorème 7.4.1 Distribution de la statistique χ^2

Lorsque H_o *est vraie, la statistique* χ^2 *suit* à peu près *une* loi χ^2_ν, *où*

$$\nu = \textit{nombre de cases} - 1$$

Notation Pour tout nombre positif $\beta < 1$, on désigne par $\chi^2_{\nu;\beta}$ le point tel que

$$P(\chi^2_\nu \geq \chi^2_{\nu;\beta}) = \beta \qquad (7.4.2)$$

où χ^2_ν est une variable aléatoire de loi khi-deux à ν degrés de liberté.

Donc la région critique est :

$$\chi^2 \geq \chi^2_{\nu;\alpha} \qquad (7.4.3)$$

où $\chi^2_{\nu;\alpha}$ est défini par (7.4.2).

Dans l'exemple, $\nu = 3$ et si on prend $\alpha =$ 0,05, le point critique est $\chi^2_{3;0,05}$ = 7,8147.

Conclusion Puisque χ^2 = 5,09 < 7,8147, on ne rejette pas H_o : on ne peut pas affirmer que le taux de suicide varie selon le jour de la semaine. L'écart entre les deux séries d'effectifs (observés et théoriques) pourrait bien être dû au hasard tout seul.

Calcul par Excel: $\chi^2_{nu;alpha}$ `=LOI.KHIDEUX.INVERSE.DROITE(alpha;nu)`.

Le modèle

Ce n'est pas tout vecteur de *k* effectifs qui se prête à un test du khi-deux. Il importe de s'assurer que le vecteur des effectifs observés est de loi multinomiale, de la statistique χ^2 en dépend.

On suppose donc que les valeurs observées constituent un vecteur

$$\boldsymbol{X} = [X_1 ; X_2 ; \ \ldots ; X_k] \sim \mathcal{M}(n ; \ p_1 ; p_2 ;\ldots ; p_k). \qquad (7.4.4)$$

L'hypothèse nulle est de la forme

$$H_o : \ p_1 = p_{1o}, \ p_2 = p_{2o}, \ldots , \ p_k = p_{ko}. \qquad (7.4.5)$$

où p_{1o}, p_{2o}, ... , p_{ko} sont des nombres positifs tels que $\sum_i p_{io} = 1$.

Les effectifs théoriques T_i sont les *espérances* des X_i sous H_o, soit

$$T_i = E(X_i \mid H_o) = np_{io} \qquad (7.4.6)$$

La statistique χ^2 est

$$\chi^2 = \sum_{i=1}^{k} \frac{(X_i - np_{io})^2}{np_{io}} \qquad (7.4.7)$$

Dans ces conditions, si l'hypothèse nulle est vraie, alors χ^2 suit à peu près une loi χ^2_ν où $\nu = k - 1$.

Remarque *Lorsque* $k = 2$, *le test d'ajustement se réduit à un test concernant une proportion p. La loi multinomiale est alors une loi binomiale, l'hypothèse* $H_o : p_1 = p_{1o}, p_2 = p_{2o}$ *est l'hypothèse* $H_o : p = p_o$, *l'énoncé* $p_2 = p_{2o}$ *(ou* $1 - p = 1 - p_o$*), étant redondant.* ■

Remarque *L'hypothèse d'une loi multinomiale*

Il faut savoir que les deux conditions qui justifient la loi multinomiale — indépendance et probabilités fixes — sont facilement violées. Dans l'exemple des suicides, deux observations qui font partie d'un pacte de suicide ne sont pas indépendantes. Dans le même exemple, si l'échantillon de 780 suicides est constitué de quatre échantillons provenant de quatre pays différents (ou de quatre époques différentes), l'hypothèse que les probabilités p_i *sont les mêmes à chaque tirage n'est pas nécessairement vérifiée.*

Un autre exemple. *Supposons qu'on tire n bulletins de votes dans une élection contestée par quatre partis. Soit* $\boldsymbol{X} = [X_1 ; X_2 ; X_3 ; X_4]$ *la distribution des votes* (X_i = *nombre de votes pour le parti* i). *Dans quelles conditions ces variables sont-elles de loi multinomiale ? Voici quelques situations où l'hypothèse d'indépendance est douteuse*:

- *Les résultats de* $n = 200$ *bulletins tirés dans une population de* 300 *bulletins ne sont pas indépendants* (*à moins que les tirages se fassent avec remise, ou que la population comprenne* 10 000 *bulletins et non* 300, *rendant les tirages pratiquement indépendants*).
- *L'hypothèse que les quatre probabilités ne changent pas d'une épreuve à l'autre n'est probablement pas vérifiée si l'échantillon est tiré de quatre circonscriptions différentes.*
- *Supposons qu'on suit la procédure suivante: on sélectionne au hasard* 10 *circonscriptions, puis on tire* 20 *bulletins dans chacune d'elles. Ici l'hypothèse d'indépendance ainsi que l'hypothèse que les probabilités sont les mêmes à chaque tirage sont toutes deux violées.*

Encore un exemple. *On tire un échantillon* (*à partir des dossiers de la police*) *de n personnes blessées dans des accidents de voiture. On classe les personnes selon le jour de la semaine où l'accident a eu lieu, ce qui donne les effectifs* $O_1, \ldots, O_7$. *L'indépendance des épreuves est-elle respectée? Certainement pas si, par exemple,* 10 *des personnes choisies ont été impliquées dans un même carambolage. Si les* O_i *comptent les accidents plutôt que les personnes impliquées, le modèle multinomial s'appliquerait sans doute bien mieux*. ■

7.5 Tests d'indépendance

Si l'hypothèse d'ajustement généralise l'hypothèse $p = p_o$, un test d'indépendance généralise le test d'égalité de deux proportions. Nous présenterons un cas particulier avant de décrire la procédure formellement. Considérons les données suivantes sur deux variables, la *scolarité S* et *l'attitude* face à l'avortement *A* :

		A : Attitude face à l'avortement			
		Pour	Mixte	Contre	
S : Scolarité	≤ 8	31	23	56	110
	9 — 12	171	89	177	437
	> 12	116	39	74	229
	Total	318	151	307	776

Ces données ont été recueillies afin de déterminer s'il y a une relation entre les deux variables. Le tableau suivant présente la distribution conditionnelle de l'attitude étant donné chaque niveau de scolarité :

		A : Attitude face à l'avortement			
		Pour	Mixte	Contre	
S : Scolarité	≤ 8	0,28	0,21	0,51	1
	9 — 12	0,39	0,20	0,41	1
	> 12	0,51	0,17	0,32	1

Une certaine dépendance se manifeste clairement dans ces distributions conditionnelles. On y voit, par exemple, que la proportion de «pour» augmente avec le niveau de scolarité. Il s'agit maintenant de savoir si cette dépendance, évidente au niveau de l'échantillon, existe aussi au niveau de la population. On doit *tester* l'hypothèse d'indépendance.

La logique du test d'indépendance est, à quelques détails techniques près, identique à celle d'un test d'ajustement : On énonce H_o, on construit un tableau d'effectifs théoriques, on mesure l'écart entre les effectifs observés et les effectifs théoriques et on rejette H_o si l'écart est trop grand.

Hypothèse nulle
L'hypothèse nulle est

H_o : les variables *scolarité* et *attitude face à l'avortement* sont indépendantes.

Calcul des effectifs théoriques
Nous devons déterminer les effectifs théoriques sous l'hypothèse d'indépendance. Ce sont les effectifs auxquels on doit s'attendre lorsque *S* et *A* sont indépendantes. Considérons la distribution marginale de la variable *Attitude face à l'avortement*, dans le tableau suivant :

		A : Attitude face à l'avortement			
		Pour	Mixte	Contre	
S: Scolarité	≤ 8				110
	9 - 12				437
	> 12				229
	Marginale	$\frac{318}{776}$	$\frac{151}{776}$	$\frac{307}{776}$	1

Si les deux variables sont indépendantes (H_o est vraie), les distributions conditionnelles de *A*, étant donné les valeurs de *S* devraient être les mêmes, et donc les mêmes que la marginale, c'est-à-dire $\frac{318}{776}$, $\frac{151}{776}$, et $\frac{307}{776}$. On répartit donc l'effectif total de chaque groupe de scolarité (110, 437, et 229) selon ces fréquences :

		A : Attitude face à l'avortement			
		Pour	Mixte	Contre	
S : Scolarité	≤ 8	$\frac{318}{776}\times 110 = 45{,}08$	$\frac{151}{776}\times 110 = 21{,}40$	$\frac{307}{776}\times 110 = 53{,}52$	110
	9 — 12	$\frac{318}{776}\times 437 = 179{,}08$	$\frac{151}{776}\times 437 = 85{,}03$	$\frac{307}{776}\times 437 = 172{,}89$	437
	> 12	$\frac{318}{776}\times 229 = 93{,}84$	$\frac{151}{776}\times 229 = 44{,}56$	$\frac{307}{776}\times 229 = 90{,}60$	229
	Total	$\frac{318}{776}\times 776$	$\frac{151}{776}\times 776$	$\frac{307}{776}\times 776$	776

Mesure de l'écart entre les effectifs observés et les effectifs théoriques
On mesure l'écart entre les effectifs observés et les effectifs théoriques exactement comme avec un test d'ajustement, soit

$$Q = \sum_{i=1}^{l} \sum_{j=1}^{c} \frac{(O_{ij} - T_{ij})^2}{T_{ij}} \tag{7.5.1}$$

Voici les calculs :

$$\begin{aligned}
\chi^2 \quad = \; & \frac{(31-45{,}08)^2}{45{,}08} + \frac{(23-21{,}40)^2}{21{,}40} + \frac{(56-53{,}52)^2}{53{,}52} \\
& + \frac{(171-179{,}08)^2}{179{,}08} + \frac{(89-85{,}03)^2}{85{,}03} + \frac{(177-172{,}89)^2}{172{,}89} \\
& + \frac{(116-93{,}84)^2}{93{,}84} + \frac{(39-44{,}56)^2}{44{,}56} + \frac{(74-90{,}60)^2}{90{,}60} \\
& = 17{,}71
\end{aligned}$$

Le point critique

On rejettera H_o si $\chi^2 \geq C$, où C est choisi de telle sorte que $P(\chi^2 \geq C) = \alpha$. Le théorème suivant permet de déterminer C.

Théorème 7.5.1 *Si les effectifs théoriques sont assez grands, la statistique* (7.5.1) *suit à peu près une loi* χ^2_ν *sous* H_o*. Le nombre de degrés de liberté est*

$$\nu = (\ell - 1) \times (c - 1)$$

où ℓ *est le nombre de lignes du tableau et* c *'est le nombre de colonnes.*

Le point critique est donc $\chi^2_{(\ell-1)(c-1)}$ et la région critique est

$$\chi^2 > \chi^2_{(\ell-1)(c-1);\alpha}$$

Ici $(\ell-1)(c-1) = (3-1)(3-1) = 4$ et $\chi^2_{4;0,05} = 9{,}4877$. Puisque $\chi^2 = 17{,}71 \geq 9{,}4877$, nous rejetons H_o à 5 %. Nous concluons que dans la population il y a vraiment une dépendance entre la scolarité et l'attitude face à l'avortement.

Le modèle

Dans un tableau de dimension $\ell \times c$, les effectifs théoriques

					Sommes
	X_{11}	X_{12}	$\ldots$	X_{1c}	$X_{1\bullet}$
	X_{21}	X_{22}	$\ldots$	X_{2c}	$X_{2\bullet}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$X_{\ell 1}$	$X_{\ell 2}$	$\ldots$	$X_{\ell c}$	$X_{\ell\bullet}$
Sommes	$X_{\bullet 1}$	$X_{\bullet 2}$	$\ldots$	$X_{\bullet c}$	n

suivent conjointement une loi multinomiale. Dans l'exemple, les X_{ij} sont les effectifs du tableau original. Les paramètres de la loi multinomiale sont n, le nombre total d'observations, et l'ensemble des probabilités

					Sommes
	p_{11}	p_{12}	$\ldots$	p_{1c}	$p_{1\bullet}$
	p_{21}	p_{22}	$\ldots$	p_{2c}	$p_{2\bullet}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$p_{\ell 1}$	$p_{\ell 2}$	$\ldots$	$p_{\ell c}$	$p_{\ell\bullet}$
Sommes	$p_{\bullet 1}$	$p_{\bullet 2}$	$\ldots$	$p_{\bullet c}$	n

Le tableau suivant présente les 9 paramètres relatifs à l'exemple, ainsi que certaines fonctions de ces probabilités, soit les sommes des lignes $p_{1\bullet}$, $p_{2\bullet}$, $p_{3\bullet}$ et les sommes des colonnes $p_{\bullet 1}$, $p_{\bullet 2}$, $p_{\bullet 3}$:

		A : Attitude face à l'avortement			
		Pour	Mixte	Contre	
S : Scolarité	≤ 8	p_{11}	p_{12}	p_{13}	$p_{1\bullet}$
	9 — 12	p_{21}	p_{22}	p_{23}	$p_{2\bullet}$
	> 12	p_{31}	p_{32}	p_{33}	$p_{3\bullet}$
	Total	$p_{\bullet 1}$	$p_{\bullet 2}$	$p_{\bullet 3}$	1

L'hypothèse nulle, en fonction de ces paramètres, est

$$H_o : \; p_{ij} = p_{i\bullet} \times p_{\bullet j} \text{ pour tout } i, j$$

Les fréquences théoriques seraient donc

		A : Attitude face à l'avortement			
		Pour	Mixte	Contre	
S : Scolarité	≤ 8	$p_{1\bullet}\, p_{\bullet 1}$	$p_{1\bullet}\, p_{\bullet 2}$	$p_{1\bullet}\, p_{\bullet 3}$	$p_{1\bullet}$
	9 — 12	$p_{2\bullet}\, p_{\bullet 1}$	$p_{2\bullet}\, p_{\bullet 2}$	$p_{2\bullet}\, p_{\bullet 3}$	$p_{2\bullet}$
	> 12	$p_{3\bullet}\, p_{\bullet 1}$	$p_{3\bullet}\, p_{\bullet 2}$	$p_{3\bullet}\, p_{\bullet 3}$	$p_{3\bullet}$
	Total	$p_{\bullet 1}$	$p_{\bullet 2}$	$p_{\bullet 3}$	1

Puisque, sous H$_o$, E(X_{ij}) = $np_{i\bullet}p_{\bullet j}$, la statistique χ^2 devrait normalement être $\sum_i \sum_j \frac{(X_{ij} - np_{i\bullet}p_{\bullet j})^2}{np_{i\bullet}p_{\bullet j}}$. Mais étant donné que les $p_{i\bullet}$ et les $p_{\bullet j}$ ne sont pas connus, elles seront estimées par

$$\hat{p}_{i\bullet} = \frac{1}{n}\sum_j X_{ij} \qquad \hat{p}_{\bullet j} = \frac{1}{n}\sum_i X_{ij}$$

La statistique χ^2 est donc

$$\chi^2 = \sum_{i=1}^{l}\sum_{j=1}^{c} \frac{(X_{ij} - n\hat{p}_{i\bullet}\hat{p}_{\bullet j})^2}{n\hat{p}_{i\bullet}\hat{p}_{\bullet j}}.$$

Étant donné le théorème 7.4.2, la région critique est

$$\chi^2 > \chi^2_{(\ell-1)(c-1);\alpha}$$

Dans l'exemple, puisque $\chi^2 = 17{,}7 > \chi^2_{4;\,0,05} = 9{,}4877$, nous rejetons H$_o$ à 5%. Nous concluons qu'il y a vraiment une dépendance entre la scolarité et l'attitude face à l'avortement.

Remarque *La notion d'indépendance*

En probabilités, l'indépendance entre deux événements A et B *est définie par* P(A∩B) = P(A)P(B). *C'est encore le sens dans lequel nous l'entendons ici. Par exemple, si* A *est l'événement «la scolarité du répondant est* ≤ 8*» et* B *l'événement «le répondant est* pour *l'avortement», alors, dans la notation du dernier tableau,* P(A∩B) = p_{11}, P(A) = $p_{1\bullet}$ et P(B) = $p_{\bullet 1}$. L*'hypothèse d'indépendance stipule que* $p_{11} = p_{1\bullet} \times p_{\bullet 1}$. ■

Remarque *Une autre modélisation*

Dans l'exemple de cette section, le chercheur a choisi 776 *sujets qu'il a ensuite classés selon la scolarité et l'attitude. Par conséquent, tous les effectifs observés sont aléatoires, y compris les effectifs des marges; seul n* = 776 *est fixe. Mais il existe des situations où les effectifs des marges sont fixés. Dans ce cas, la description des données comme un vecteur de loi multinomiale n'est pas conforme à la réalité. Le tableau suivant présente des données recueillies dans le but de savoir si la durée d'une hospitalisation* (pour une certaine maladie) *varie d'un hôpital à l'autre* :

		Durée de l'hospitalisation			
		1-2 *jours*	3-5 *jours*	*plus de* 5 *jours*	
Hôpital	1	40	20	40	100
	2	50	35	65	150
	3	95	45	60	200

Visiblement, les effectifs de la marge de droite sont fixes : on a décidé, avant de prélever les données, qu'on prendrait respectivement 100, 150 *et* 200 *cas dans les trois hôpitaux. Ce ne sont pas des variables aléatoires.*

Dans le tableau suivant on nomme les variables observées :

		Durée de l'hospitalisation			
		1-2 jours	3-5 jours	5 jours+	
Hôpital	1	X_1	Y_1	Z_1	n_1
	2	X_2	Y_2	Z_2	n_2
	3	X_3	Y_3	Z_3	n_3

Les 9 variables aléatoires dans le tableau ne suivent pas une loi multinomiale. Nos observations sont constituées de 3 vecteurs de loi multinomiale. Le modèle est plutôt celui-ci :

$$[X_1 ; Y_1 ; Z_1] \sim \mathcal{M}(n_1 ;\ p_1 ;\ r_1 ;\ s_1) ;$$
$$[X_2 ; Y_2 ; Z_2] \sim \mathcal{M}(n_2 ;\ p_2 ;\ r_2 ;\ s_2) ;$$
$$[X_3 ; Y_3 ; Z_3] \sim \mathcal{M}(n_3 ;\ p_3 ;\ r_3 ;\ s_3).$$

L'hypothèse nulle est

$$\mathrm{H_o} :\ [p_1 ; r_1 ;\ s_1] = [p_2 ; r_2 ; s_2] = [p_3 ; r_3 ;\ s_3] \qquad \blacksquare$$

Malgré la différence entre ce modèle (*trois multinomiales*) *et celui de l'exemple précédent* (*une seule multinomiale*), *la procédure reste la même : le même test est valable dans les deux cas.*

Remarque *Le cas particulier d'un tableau* 2×2 *est un test d'égalité de deux proportions. La formulation habituelle du test d'égalité de proportions est celle-ci : on observe deux variables aléatoires indépendantes,* $X_1 \sim \mathcal{B}(n_1 ; p_1)$ *et* $X_2 \sim \mathcal{B}(n_2 ; p_2)$, *et on teste l'hypothèse* $\mathrm{H_o} : p_1 = p_2$, *un test essentiellement équivalent à un test d'indépendance. Mais le déroulement de l'expérience ne mène pas toujours à cette modélisation. Par exemple, supposons qu'on veuille tester l'hypothèse que la proportion des personnes favorables à une certaine loi (permettant le mariage gai, par exemple) est la même chez les femmes et les hommes. Les données prennent la forme suivante :*

	Pour projet de loi	Contre le projet de loi	Total
Femmes	X_1	$Y_1 = n_1 - X_1$	$n_1 = X_1 + Y_1$
Hommes	X_2	$Y_2 = n_2 - X_2$	$n_2 = X_2 + Y_2$
	$X_1 + X_2$	$Y_1 + Y_2$	$n = n_1 + n_2$

Si le nombre de femmes n_1 *et le nombre d'hommes* n_2 *sont fixes, alors effectivement le modèle de deux binomiales indépendantes s'impose directement. Mais il est possible que le tirage soit fait au hasard dans la population entière. Dans ce cas, nos observations constituent un vecteur* $[X_1 ; Y_1 ; X_2 ; Y_2]$ *de loi* $\mathcal{M}(n ; p_{11} ; p_{12} ; p_{21} ; p_{22})$. *La raison pour laquelle on peut néanmoins continuer à agir comme si on avait affaire à deux variables indépendantes de loi binomiale, c'est que les variables* X_1 *et* X_2 *sont* conditionnellement *indépendantes et de lois binomiales. C'est-à-dire, étant donné* $X_1 + Y_1 = n_1$ *et* $X_2 + Y_2 = n_2$:

$$X_1|X_1 + Y_1 = n_1 \sim \mathcal{B}(n_1 ; p_1) \text{ et } X_2|X_2 + Y_2 = n_2 \sim \mathcal{B}(n_2 ; p_2), \text{ où}$$

$$p_1 = \frac{p_{11}}{p_{11} + p_{12}} \text{ et } p_2 = \frac{p_{21}}{p_{21} + p_{22}} . \qquad \blacksquare$$

7.6 Données appariées (test de McNemar)

Le conditionnement décrit dans la dernière remarque est utile aussi dans un autre contexte, le test de McNemar. Il s'agit d'un test d'égalité de proportions lorsque les données sont appariées. Cette situation est tout à fait parallèle au test d'égalité de *moyennes* dans le cas de données appariées. L'hypothèse à tester est $p_1 = p_2$, mais contrairement au cas présenté à la section 7.3, les estimateurs $\hat{p}_1$ et $\hat{p}_2$ ne sont pas indépendants. La solution, illustrée dans l'exemple suivant, est basée sur une distribution *conditionnelle*.

Supposons qu'on doive évaluer l'efficacité d'une pilule conçue pour réduire les risques d'une crise de panique chez les personnes qui en sont susceptibles. Une façon de procéder est la suivante : on soumet deux groupes — un groupe expérimental qui reçoit le traitement et un groupe témoin qui reçoit un placebo — à une situation de stress susceptible de provoquer une crise. Les variables X_1 et X_2 (nombre de réactions de panique enregistrées dans les deux groupes), sont indépendantes, de loi binomiale de probabilités p_1 et p_2 respectivement. Les techniques de la section 7.3 peuvent alors être utilisées pour tester l'hypothèse $p_1 = p_2$.

Mais un autre plan d'expérience consiste à placer chaque individu dans une situation de stress *deux fois*, l'une avec le traitement, l'autre avec le placebo. On observe encore les variables X_1 et X_2, elles sont de loi binomiale de paramètres p_1 et p_2, et c'est encore l'hypothèse $p_1 = p_2$ qu'il s'agit de tester. Mais on ne peut pas procéder comme d'habitude parce que X_1 et X_2 ne sont pas indépendantes. Nous observons en fait un vecteur de quatre variables aléatoires, disons $\boldsymbol{X} = [X_{11} ; X_{12} ; X_{21} ; X_{22}]$, de loi multinomiale de paramètres n (le nombre de sujets) et p_{11}, p_{12}, p_{21} et p_{22}, tel que présentés dans la tableau suivant :

		Sous traitement		
		Réaction	Pas de réaction	
Sous placebo	Réaction	X_{11} p_{11}	X_{12} p_{12}	$X_1 = (X_{11} + X_{12})$ p_1
	Pas de réaction	X_{21} p_{21}	X_{22} p_{22}	$n - X_1$
		$X_2 = (X_{11} + X_{21})$ p_2	$n - X_2$	n

L'hypothèse à tester est

$$\mathrm{H_o :}\ \ p_1 = p_2 \Leftrightarrow p_{11} + p_{12} = p_{11} + p_{21} \Leftrightarrow p_{12} = p_{21}$$

On peut tester cette hypothèse en considérant la distribution *conditionnelle* de X_{12} étant donné la valeur observée de $X_{12} + X_{21}$:

$$X_{12} \big| \ X_{12} + X_{21} = m \ \sim \mathcal{B}\left(m \ ; \ \frac{p_{12}}{p_{12} + p_{21}} \right)$$

$\mathrm{H_o}$ est équivalente à $\mathrm{H_o}$: $\dfrac{p_{12}}{p_{12} + p_{21}} = 1/2$. Il s'agit donc d'un test classique sur une proportion.

Exemple 7.6.1 *Test d'égalité de proportions avec données appariées*

Considérons donc les données suivantes sur un échantillon de 150 personnes soumises deux fois à une situation de stress, l'une sous l'effet d'un médicament, l'autre sous l'effet d'un placebo.

		Sous traitement		
		Réaction de panique	*Pas de réaction de panique*	*Total*
Sous placebo	*Réaction de panique*	12	28	40
	Pas de réaction de panique	13	47	60
		25	75	100

On considère les 41 (13+28) sujets dont le comportement diffère sous le traitement et sous le placebo. On teste l'hypothèse que parmi ceux-ci, la probabilité que la réaction de panique se produise sous l'effet du traitement est 1/2. La variable observée X suit conditionnellement une loi

$\mathcal{B}(41\,;p)$, H$_o$: $p = ½$. $\hat{p} = \dfrac{28}{41}$ et $Z = \dfrac{28/41 - 1/2}{\frac{1}{2}\sqrt{\frac{1}{41}}}$ = 2,34, ce qui est hautement significatif, le seuil expérimental (pour un test bilatéral) étant 0,019. Parmi ceux dont la réponse diffère sous l'effet du médicament, ceux qui ont mieux réagi sous l'effet du médicament que sous l'effet du placebo sont plus nombreux (28/41) que les autres. On conclut que le médicament est bénéfique. ■

7.7 Sommes de khi-deux

Considérons le tableau suivant, qui donne la distribution des quotients intellectuels (QI) d'un échantillon d'étudiants pris dans chacune de quatre universités :

	Université				
	A	B	C	D	Toutes
QI < 115	12	30	30	46	118
115 ≤ QI <125	47	49	34	18	148
QI ≥125	41	21	36	36	134
Effectif total	100	100	100	100	400

Les quatre colonnes sont des vecteurs

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix};\ \boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix};\ \boldsymbol{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix};\boldsymbol{W} = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix},$$

chacun de loi multinomiale de paramètres n = 100 et

$$\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix};\ \boldsymbol{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix};\ \boldsymbol{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix};\ \text{et } \boldsymbol{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix},$$

respectivement. L'hypothèse d'indépendance stipule que ces quatre vecteurs sont égaux ($\boldsymbol{p} = \boldsymbol{q} = \boldsymbol{r} = \boldsymbol{s}$), mais ne spécifie par leur valeur. Il peut arriver, cependant, qu'on veuille tester une hypothèse de la forme :

$$\boldsymbol{p} = \boldsymbol{q} = \boldsymbol{r} = \boldsymbol{s} = \boldsymbol{p}_o$$

où $\boldsymbol{p}_o$ est un vecteur donné. Par exemple,

$$\text{H}_o : \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

(Pourquoi 1/3? Ces probabilités correspondent à l'hypothèse que, dans chaque université, les QI sont de loi normale de moyenne 120 et d'écart-type 12, ces paramètres étant une norme validée au cours des années sur un très grand nombre d'individus, d'âge et d'éducation comparables à ceux des étudiants de ces universités).

On peut tester cette hypothèse en effectuant un test d'uniformité dans chaque université séparément. On obtient les résultats suivants :

	Université			
	A	B	C	D
χ^2	21,02	35,06	0,56	12,08
vp	0,00003	2×10^{-8}	0,756	0,0024

Chacune des valeurs de χ^2 est de loi L'hypothèse d'uniformité est acceptée à l'université C et rejetée aux trois autres. Mais là nous nous retrouvons avec quatre conclusions alors que nous n'avons qu'une seule hypothèse :

$$\mathrm{H_o} : \boldsymbol{p} = \boldsymbol{p}_o \text{ et } \boldsymbol{q} = \boldsymbol{p}_o \text{ et } \boldsymbol{r} = \boldsymbol{p}_o \text{ et } \boldsymbol{s} = \boldsymbol{p}_o, \text{ où } \boldsymbol{p}_o = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Rejeter cette hypothèse, c'est conclure qu'au moins l'une de ces affirmations est fausse. Une statistique appropriée est simplement la somme des valeurs de χ^2 :

$$\chi^2 = 21{,}02 + 35{,}06 + 0{,}56 + 12{,}08 = 68{,}72.$$

Les quatre échantillons sont indépendants et sous $\mathrm{H_o}$ chacune des valeurs de χ^2 est de loi χ^2 (à peu près) à 2 degrés de liberté. Leur somme est donc de loi χ^2 à 8 degrés de liberté. La valeur observée $\chi^2 = 68{,}72$ correspond à une *valeur p* de $vp = 8{,}8\times10^{-12}$. Donc on rejette l'hypothèse, pour conclure qu'il n'est pas vrai que la distribution des QI est normale (de moyenne 120 et d'écart-type 12) dans chacune des universités.

Remarque *On pourrait songer à l'approche suivante. On combine les données des quatre universités en une seule série, soit* $\boldsymbol{U} = \boldsymbol{X} + \boldsymbol{Y} + \boldsymbol{Z} + \boldsymbol{W}$ = [118; 148; 134]. *Sous* $\mathrm{H_o}$, *ces effectifs sont de loi multinomiale de paramètres n* = 400 *et probabilités* [1/3 ; 1/3 ; 1/3]. *Un test d'ajustement donne une valeur de* χ2 *de* 3,38 *à* 2 *degrés de liberté. Soit vp* = 0,18, *on ne peut pas rejeter* $\mathrm{H_o}$.

Le premier rejette $\mathrm{H_o}$, *alors que le deuxième, basé sur* $\boldsymbol{U}$, *ne rejette pas* $\mathrm{H_o}$. *Auquel faut-il croire?*

Strictement parlant, les deux tests sont corrects dans le sens suivant. Dans les deux cas, la probabilité d'une erreur de première espèce est d'environ 5 %, *et dans ce sens, les deux tests sont corrects. C'est au niveau de la puissance que les deux tests diffèrent—la capacité du test de rejeter* $\mathrm{H_o}$ *lorsqu'elle est fausse. Le deuxième test, qui repose entièrement sur le vecteur* $\boldsymbol{U}$, *est insensible aux différences entre* $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ et $\boldsymbol{W}$, *et donc ne contient aucune information sur les différences entre* $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$, *et* $\boldsymbol{s}$. *Il accepte* $\mathrm{H_o}$ *parce que les composantes de* $\boldsymbol{U}$ *sont assez proches les unes des autres. Or il y a d'importantes différences entre* $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ *et* $\boldsymbol{W}$. *Le deuxième test détecte ces différences et conclut à des différences entre* $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$, *et* $\boldsymbol{s}$. ■

7.8 Comparaisons multiples

Dans certaines études scientifiques, une erreur de première espèce constitue ce qu'on appelle une *fausse découverte*. Une question courante en génétique, par exemple, est de savoir si un gène donné est lié à une certaine maladie. $\mathrm{H_o}$ est l'hypothèse que le gène n'est pas lié à la maladie. Si on rejette $\mathrm{H_o}$ on affirme avoir fait une découverte. Si l'hypothèse est vraie, c'est une *fausse découverte*. La probabilité d'une fausse découverte est α, le niveau du test. On se protège donc contre une fausse découverte en prenant α petit.

Une difficulté surgit, cependant, lorsqu'un laboratoire, à la recherche d'un gène pouvant être lié à une même maladie considère un grand nombre *n* de gènes possibles, donc *n* tests d'hypothèses. Si le risque d'une fausse découverte est maintenu à α à chacun des tests, le risque d'au moins une fausse découverte au cours des *n* tests croît avec *n*. Plus précisément, si *n* est assez grand, il est très probable de rejeter au moins une des hypothèses, mêmes si elles sont toutes vraies. Le défi est donc celui-ci:

Lorsqu'on teste *n* hypothèses, comment contrôler la probabilité de rejeter au moins l'une d'elles si elles sont toutes vraies?

La méthode de Bonferroni

Soit $\mathrm{H_{o1}}, \mathrm{H_{o2}}, \ldots, \mathrm{H}_{on}$ *n* hypothèses, et $\mathrm{H_o}$ la conjonction des *n* hypothèses $\mathrm{H_{o1}}, \mathrm{H_{o2}}, \ldots, \mathrm{H}_{on}$. On désigne $\mathrm{H_o}$ par

$$\mathrm{H_o} = \mathrm{H_{o1}} \wedge \mathrm{H_{o2}} \wedge \ldots \wedge \mathrm{H}_{on}.$$

($\mathrm{H_o}$ est vraie si les *n* hypothèses sont toutes vraies).

Soit $R_1, R_2, \ldots, R_n$ n régions critiques correspondant aux n hypothèses, et R_o une région critique pour H_o. Le problème est de s'assurer que $P(R_i \mid H_{oi})$ assez petit pour que $P(R_o \mid H_o) \leq \alpha$.

On vérifie aisément que si les R_i sont tels que $P(R_i \mid H_{oi}) \leq \alpha/n$ pour tout i, et

$$R_o = R_1 \cup R_2 \cup \ldots \cup R_n,$$

alors $P(R_o \mid H_o) \leq \alpha$. Ceci découle du fait que la probabilité de l'union de n événements est inférieure ou égale à la somme de leurs probabilités:

$$P(R_o \mid H_o) = P(R_1 \cup R_2 \cup \ldots \cup R_n) \leq P(R_1) + P(R_2) + \ldots + P(R_n).$$

Si les hypothèses sont toutes vraies, chacun des événements R_i est de probabilité α/n (ou inférieures à α/n), et par conséquent, $P(R_1) + P(R_2) + \ldots + P(R_n) \leq n(\alpha/n) = \alpha$.

La méthode de Bonferroni consiste donc à effectuer chaque test au niveau α/n.

Exemple 7.8.1 *Test simultané sur plusieurs proportions*

Les chercheurs d'un laboratoire s'intéressent aux causes génétiques d'une certaine maladie. Certaines considérations théoriques leur fait penser que 4 gènes particuliers (les gènes g_1, g_2, g_3 et g_4) pourraient être en cause. Afin de confirmer cette hypothèse, on rassemble deux groupes de personnes. Le groupe 1 (le groupe «expérimental») est constitué de n_1 = 70 sujets souffrant de la maladie en question. Le groupe 2 (le groupe témoin), de n_2 = 200 personnes n'en est pas atteint. On compte alors le nombre de personnes—X_i dans le groupe 1 et Y_i dans le groupe 2—qui portent le gène g_i.

$$X_i \sim \mathcal{B}(70\,;\,p_{1i}) \text{ et } Y_i \sim \mathcal{B}(200\,;\,p_{2i})\,.$$

Si on conclut que $p_{1i} > p_{2i}$, on est en mesure d'affirmer (du moins provisoirement) que le gène g_i est en partie responsable de la maladie.

On a 4 hypothèses à tester:

$$H_{oi} : p_{1i} = p_{2i}, \text{ contre } H_{1i} : p_{1i} > p_{2i}\,,\ i = 1, 2, 3, 4.$$

(Le choix d'une alternative unilatérale pourrait, dépendant du contexte, être contesté, mais nous le supposons justifiable ici).

Voici les résultats:

Nombre de sujets possédant le gène	Gène (g_i)			
	g_1	g_2	g_3	g_4
Dans le groupe 1 (X_i)	17	22	23	35
Dans le groupe 2 (Y_i)	25	38	59	77

Posons α = 0,05. On effectue chaque test au niveau $\alpha/4$ = 0,0125.
Les statistiques de test sont

$$Z_i = \frac{\hat{p}_{1i} - \hat{p}_{21}}{\hat{p}_i(1-\hat{p}_i)\sqrt{\dfrac{1}{70}+\dfrac{1}{200}}}\,, \text{ où } \hat{p}_i = \frac{70\,\hat{p}_{1i} + 200\,\hat{p}_{2i}}{70+200}\,.$$

Les régions critiques sont $Z_i \geq z_{0,0125} = 2,24$.

Voici les calculs:

	Gène g_1	Gène g_2	Gène g_3	Gène g_4
$\hat{p}_{1i}$	0,2429	0,3143	0,3286	0,5
$\hat{p}_{2i}$	0,1250	0,1900	0,2950	0,3850
$\hat{p}_{i}$	0,1556	0,2222	0,3037	0,4148
z_i	2,3416	2,1527	0,5257	1,6808
vp	0,0096	0,0157	0,2996	0,0464

Puisque $z_1 > 2{,}24$, on peut rejeter les hypothèses H_{o1} et conclure donc que $p_{11} > p_{21}$: on affirme que le gène g_1 est lié à la maladie. On ne peut le dire d'aucun des autres gènes.

Remarque Le but de cet exemple est d'illustrer la méthode de Bonferroni. L'analyse qui y est faite est possiblement simpliste dans la mesure où elle ne tient pas compte des interactions possibles entre les gènes.

Remarque *Au dernier exemple, les régions critiques sont exprimées en fonction des valeurs Z_i. Il est souvent plus commode de s'exprimer en fonction des valeurs p. La région critique $\{Z \geq z_{0,0125}\}$ est équivalente à :* $\{vp > 0{,}0125\}$.

Procédure de Hochberg

[Yosef Hochberg, A sharper Bonferroni procedure for multiple tests of significance, *Biometrika* (1988), **75**, 4, pp. 800-2]

Un des inconvénients de la méthode de Bonferroni est qu'elle oblige à réduire la taille de la région critique à α/n, ce qui a pour conséquence que le rejet d'une hypothèse devient très difficile si n est grand. Certains proposent de choisir α plus grand—jusqu'à 30 %, s'il le faut½–, donc de s'exposer à un risque plus élevé d'annoncer des découvertes qui n'en sont pas.

Mais il existe une amélioration de la procédure de Bonferroni, la procédure de Hochberg, qui qui réduit quelque peu le risque de fausses découvertes. On commence par exprimer les régions critiques en termes des *valeurs p*, soit $\{p \leq \alpha\}$ pour un test d'hypothèse de niveau α.

Description de la procédure:

Soit $p_1, p_2, \ldots, p_n$ les *valeurs p* correspondant aux hypothèses $H_{o1}, H_{o2}, \ldots, H_{on}$; $p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(n)}$ les statistiques d'ordre et $H_{(o1)}, H_{(o2)}, \ldots, H_{(on)}$ les hypothèses correspondantes.

On calcule les quantités

$$r_1 = np_{(1)} \ ; \ r_2 = (n-1)p_{(2)} \ ; \ r_2 = (n-2)p_{(3)} \ ; \ \ldots \ ; \ r_n = p_{(n)} \ ;$$

On rejette toute hypothèse $H_{(oi)}$ pour laquelle $r_i \leq \alpha$.

Justification de la procédure de Hochberg

Par comparaison, la méthode de Bonferroni rejette toute hypothèse $H_{(oi)}$ pour laquelle $np_{(i)} \leq \alpha$. Il est donc possible que l'approche de Hochberg permette de rejeter une hypothèse qui ne pourrait être rejetée par la méthode de Bonferroni. C'est le cas de l'hypothèse $H_{(3)}$, par exemple, si $(n-2)p_{(3)} \leq \alpha < np_{(3)}$.

Cette méthode possède la propriété voulue, soit que $P(R_o \mid H_o) \leq \alpha$, et elle répond aussi à un souci de cohérence. Par exemple, si on teste deux hypothèses, soit H_{o1} et H_{o2}, ainsi que $H_o = H_{o1} \wedge H_{o2}$, il serait incohérent de rejeter H_{o1} et d'accepter H_o. Il faudrait donc ne rejeter H_{o1} que si on rejette également H_o. Si R_1, R_2 et $R_o = R_1 \cup R_2$ sont des régions critiques, chacune de taille α, alors pour éviter toute incohérence, on définit de nouvelles régions critiques $R_1^* = R_1 \cap R_o$, $R_2^* = R_2 \cap R_o$ et

$R_o^* = R_1^* \cup R_2^*$. Puisque $R_1^* \subset R_o$ et $R_2^* \subset R_o$ et $R_o^* \subset R_o$ sous H_o, $P(R_1^* \mid H_o) \leq \alpha$, $P(R_2^* \mid H_o) \leq \alpha$ et $P(R_o^* \mid H_o) \leq \alpha$.

En général, la procédure se définit comme suit.

Considérons les hypothèses H_{o1}, H_{o2}, … , H_{on}, ainsi que toutes les conjonctions possibles, telles, par exemple, $H_{o1} \wedge H_{o2}$, $H_{o2} \wedge H_{o3}$, $H_{o1} \wedge H_{o2} \wedge H_{o3}$, $H_{o1} \wedge H_{o2} \wedge \ldots \wedge H_{on}$. Supposons qu'il existe une région critique de niveau α pour chacune de ces hypothèses, et que dans le cas d'une conjonction de 2 hypothèses ou plus, la région critique est définie par l'approche de Bonferroni. Le test de Hochberg consiste à rejeter H_{oi} si et seulement si on rejette H_{oi} ainsi que toute conjonction d'hypothèses qui comprend H_{oi}. Cette procédure mène à la procédure de Hochberg telle que décrite ci-dessus.

On justifie la procédure cas $n = 3$. Soit $p_{(1)} \leq p_{(2)} \leq p_{(3)}$ les *valeurs p* ordonnées et $H_{(o1)}$, $H_{(o2)}$, $H_{(o3)}$ les hypothèses correspondantes. Voici les conditions dans lesquelles on rejette $H_{(o1)}$:

$H_{(o1)}$ et toutes les conjonctions la contenant	Région critique (selon Bonferroni)	Expression simplifiée
$H_{(o1)}$	$p_{(1)} \leq \alpha$	$p_{(1)} \leq \alpha$
$H_{(o1)} \wedge H_{(o2)}$	$p_{(1)} \leq \alpha/2$ ou $p_{(2)} \leq \alpha/2$	$p_{(1)} \leq \alpha/2$
$H_{(o1)} \wedge H_{(o3)}$	$p_{(1)} \leq \alpha/2$ ou $p_{(3)} \leq \alpha/2$	$p_{(1)} \leq \alpha/2$
$H_{(o1)} \wedge H_{(o2)} \wedge H_{(o3)}$	$p_{(1)} \leq \alpha/3$ ou $p_{(2)} \leq \alpha/3$ ou $p_{(3)} \leq \alpha/3$	$p_{(1)} \leq \alpha/3$

On rejette $H_{(o1)}$ si $p_{(1)} \leq \alpha$ *et* $p_{(1)} \leq \alpha/2$ *et* $p_{(1)} \leq \alpha/3$. Donc on rejette $H_{(o1)}$ si $p_{(1)} \leq \alpha/3$.

Voici les conditions dans lesquelles on rejette $H_{(o2)}$:

$H_{(o1)}$ et toutes les conjonctions la contenant	Région critique (selon Bonferroni)	Expression simplifiée
$H_{(o2)}$	$p_{(2)} \leq \alpha$	$p_{(2)} \leq \alpha$
$H_{(o1)} \wedge H_{(o2)}$	$p_{(1)} \leq \alpha/2$ ou $p_{(2)} \leq \alpha/2$	$p_{(1)} \leq \alpha/2$
$H_{(o2)} \wedge H_{(o3)}$	$p_{(2)} \leq \alpha/2$ ou $p_{(3)} \leq \alpha/2$	$p_{(2)} \leq \alpha/2$
$H_{(o1)} \wedge H_{(o2)} \wedge H_{(o3)}$	$p_{(1)} \leq \alpha/3$ ou $p_{(2)} \leq \alpha/3$ ou $p_{(3)} \leq \alpha/3$	$p_{(1)} \leq \alpha/3$

On rejette $H_{(o2)}$ si $p_{(2)} \leq \alpha$ *et* $p_{(1)} \leq \alpha/2$ *et* $p_{(2)} \leq \alpha/2$ *et* $p_{(1)} \leq \alpha/3$, donc si $p_{(1)} \leq \alpha/3$ et $p_{(2)} \leq \alpha/2$.

Voici les conditions dans lesquelles on rejette $H_{(o3)}$:

$H_{(o3)}$ et toutes les conjonctions la contenant	Région critique (selon Bonferroni)	Expression simplifiée
$H_{(o3)}$	$p_{(3)} \leq \alpha$	$p_{(3)} \leq \alpha$
$H_{(o1)} \wedge H_{(o3)}$	$p_{(1)} \leq \alpha/2$ ou $p_{(3)} \leq \alpha/2$	$p_{(1)} \leq \alpha/2$
$H_{(o2)} \wedge H_{(o3)}$	$p_{(2)} \leq \alpha/2$ ou $p_{(3)} \leq \alpha/2$	$p_{(2)} \leq \alpha/2$
$H_{(o1)} \wedge H_{(o2)} \wedge H_{(o3)}$	$p_{(1)} \leq \alpha/3$ ou $p_{(2)} \leq \alpha/3$ ou $p_{(3)} \leq \alpha/3$	$p_{(1)} \leq \alpha/3$

On rejette $H_{(o3)}$ si $p_{(3)} \leq \alpha$ *et* $p_{(1)} \leq \alpha/2$ *et* $p_{(2)} \leq \alpha/2$ *et* $p_{(1)} \leq \alpha/3$, donc si $p_{(1)} \leq \alpha/3$, $p_{(2)} \leq \alpha/2$ et $p_{(3)} \leq \alpha$.

Pour illustrer la différence entre cette méthode et la méthode de Bonferroni, considérons les conditions qui permettent de rejeter rejette $H_{(01)}$, $H_{(02)}$ et $H_{(03)}$. Selon la méthode modifiée, on rejette $H_{(01)}$, $H_{(02)}$ et $H_{(03)}$ si

$$p_{(1)} \leq \frac{\alpha}{n}\ ,\ p_{(2)} \leq \frac{\alpha}{n-1}\ \text{ et } p_{(3)} \leq \frac{\alpha}{n-2}\ ,$$

alors que par la méthode de Bonferroni ces mêmes trois hypothèses ne sont rejetées que lorsque

$$p_{(1)} \leq \frac{\alpha}{n}\ ,\ p_{(2)} \leq \frac{\alpha}{n} \text{ et } p_{(3)} \leq \frac{\alpha}{n}\,.$$

Exemple 7.8.2 *La méthode de Hotchberg*

On applique la méthode de Hochberg au contexte de l'exemple 7.8.1.

Les *valeurs p* ordonnées sont 0,010 0,016 0,046 0,300
Les valeurs $(n - i + 1)p(i)$ sont : 0,038 0,047 0,093 0,300.
Si le niveau souhaité est α = 0,05, on rejette les hypothèses $H_{(o1)}$ et $H_{(o2)}$ (qui sont aussi H_{o1} et H_{o2}). L'approche Bonferroni n'aurait pas permis de rejeter H_{o2}.

6.9 Trois problèmes résolus

Problème 1

Pour des raisons d'économie, un chocolatier compte remplacer le chocolat traditionnel T par un nouveau chocolat N (même formule que pour T, avec un ingrédient exotique en moins) et se demande si la différence est vraiment perceptible. Il fait tester les deux chocolats par un expert en lui demandant d'identifier 12 chocolats, dont certains sont traditionnels (T), d'autres nouveaux (N). Posons comme hypothèse nulle l'hypothèse

H_o : La différence entre les deux sortes est imperceptible

a) Considérons le plan suivant :

Plan A : On lui fait goûter 6 chocolats T et 6 N, sans lui dire combien il y en a de chaque.

Supposons que l'expert réussit à identifier 10 des chocolats. Peut-on rejeter H_o ?

Réponse

Si X le nombre de succès, nous allons calculer $P(X \geq 10)$ sous H_o. Dans ce plan, on suppose que les essais sont indépendants et que la probabilité de succès est ½ à chaque essai. Donc X est de loi $\mathcal{B}$(10 ; ½), et $P(X \geq 10 \mid p = ½) = 0{,}01928711$. Cette probabilité étant plutôt faible, nous rejetons H_o : le sujet semble bien avoir certaines capacités.

b) Considérons le plan suivant :

Plan B : On lui fait goûter 6 chocolats T et 6 N, mais il sait qu'il y en a 6 de chaque

Supposons que l'expert réussit à identifier 10 des chocolats. Peut-on rejeter H_o ?

Réponse

$X = 2Y$, où Y, le nombre de chocolats T identifiés comme T, est de loi hypergéométrique de paramètres $n = 6$, $N_1 = N_2 = 6$.

		Le chocolat est		
		T	N	
Le chocolat est déclaré	T	Y	6-Y	6
	N	6-Y	Y	6
		6	6	12

$P(X \geq 10) = P(Y \geq 5) = 0{,}0400$.

c) Considérons le plan suivant :

Plan C : On lui fait goûter 4 chocolats T et 8 N, et il sait qu'il y en a 4 T et 8 N.

Supposons que l'expert réussit à identifier 10 des chocolats. Peut-on rejeter H_o ?

Réponse

$X = 2Y + 4$, où Y est de loi hypergéométrique de paramètres $n = 4$, $N_1 = 4$, $N_2 = 8$.

		Le chocolat est		
		T	N	
Le chocolat est déclaré	T	Y	4-Y	4
	N	4-Y	4+Y	8
		4	8	12

P($X \geq 10$) = P($Y \geq 3$) = 0,.0667
Le test A étant le plus exigent des trois, le fait d'avoir réussi 10 fois est particulièrement significatif, ce qui l'est moins dans le cas des tests B et C, plus faciles.

Problème 2 [*Suite*]
Supposons que l'expérience décrite au dernier numéro n'a pas encore été faite et qu'on doit la planifier, c'est-à-dire, déterminer une région critique. Celle-ci sera exprimée en fonction de X, le nombre de succès, et la région critique sera de la forme $X \geq a$. On décide qu'on rejettera H_o si le nombre de succès est assez élevé, donc si X est supérieur ou égal à un certain nombre a. On choisira a de telle sorte que P($X \geq a|H_o$) ≤ α. Fixons α = 0,05 Déterminer a pour chacun des plans, prenant pour seuil α = 0,05.

Plan A

Le nombre de succès X est de loi $\mathcal{B}$(12 ; ½). Voici les valeurs de P($X \geq a$ | H_o) pour différentes valeurs de a :

a	P($X \geq a$)
12	0,0002
11	0,0032
10	0,0193
9	0,0730
8	0,1938
7	0,3872
6	0,6128

Si $X \geq 10$, on rejettera l'hypothèse. La région critique est de taille 0,0193.

Plan B

Le nombre de succès X est égal à 2Y, où Y est de loi hypergéométrique de paramètres n = 6, $N_1 = N_2$ = 6.

Nombre de succès X = 2Y	Nombre de T déclarés T (Y)	P($X \geq a$)
12	6	0,00108
10	5	0,0400
8	4	0,2835
6	3	0,7165
4	2	0,9600
2	1	0,9989
0	0	1

On ne rejettera H_o que si le sujet réussi à identifier au moins 10 des chocolats. La région critique est de taille 0,0400.

Plan C

Le nombre de succès X est égal 2Y+4, où Y est de loi hypergéométrique de paramètres n = 4, N_1 = 4, N_2 = 8.

Nombre de succès $a = 2y+4$	Nombre de T déclarés T (y)	$P(X \geq a)$
12	4	0,0020
10	3	0,0667
8	2	0,4061
6	1	0,8586
4	0	1

On ne rejettera H_o que si le sujet réussi à identifier tous les chocolats. La région critique est de taille 0,0020.

Problème 3

Une personne prétend être capable de distinguer 15 marques de bière. Pour le prouver, il se soumet à un test qui lui demande de les identifier (il prend une gorgée de chaque sorte). Considérons l'hypothèse

H_o : Cette personne est totalement incapable de distinguer une bière de l'autre.

Soit X le nombre de marques de bière correctement identifiées. Voici la distribution de X sous l'hypothèse H_o :

a	$P(X = a)$	$P(X \geq a)$	a	$P(X = a)$	$P(X \geq a)$	a	$P(X = a)$	$P(X \geq a)$
0	0,36788	1	6	0,00051	0,00059	11	0,00000	0,00000
1	0,36788	0,63212	7	0,00007	0,00008	12	0,00000	0,00000
2	0,18394	0,26424	8	0,00001	0,00001	13	0,00000	0,00000
3	0,06131	0,08030	9	0,00000	0,00000	14	0,00000	0,00000
4	0,01533	0,01899	10	0,00000	0,00000	15	0,00000	0,00000
5	0,00307	0,00366						

Si X est assez élevée on rejettera H_o.

Réponse À un seuil $\alpha = 0{,}05$, on rejettera H_o s'il identifie correctement 4 marques ou plus.

Résumé

Expression générale

La multiplicité des tests dans ce chapitre et le précédent est à première vue déconcertante. Il est bon de se rappeler que si les détails sont nombreux, les concepts sous-jacents ne le sont pas. Tous les tests traités dans ce manuel (à l'exception des tests sur les variances), suivent essentiellement la même démarche, une démarche que l'on peut décrire en des termes généraux :

- Une hypothèse H_o porte sur un paramètre θ ;
- L'hypothèse nulle est de la forme $\theta = \theta_o$;
- Le test est basé sur un estimateur $\hat{\theta}$ de θ d'écart-type $\sigma_{\hat{\theta}}$;
- Il existe un estimateur $\hat{\sigma}_{\hat{\theta}}$ de $\sigma_{\hat{\theta}}$;
- La statistique de test est une mesure normalisée de l'écart entre $\hat{\theta}$ et θ, soit $T = \dfrac{\hat{\theta} - \theta_o}{\hat{\sigma}_{\hat{\theta}}}$;
- La région critique est de la forme $|T| \geq$ A, ou $T \geq$ B, ou $T \leq$ C selon que H_1 est $\theta \neq \theta_o$, $\theta \geq \theta_o$, ou $\theta \leq \theta_o$;
- Les points critiques A, B, ou C sont choisis de telle sorte que P($|T| \geq$ A) = α, P($T \geq$ B) = α, P($T \leq$ C) = α ;
- Les points critiques A, B et C dépendent de la loi de la statistique T.

Variance des estimateurs :

Paramètre (θ)	Estimateur ($\hat{\theta}$)	Écart-type de l'estimateur $\sigma_{\hat{\theta}}$	Estimateur de $\sigma_{\hat{\theta}}$ ($\hat{\sigma}_{\hat{\theta}}$)
μ	$\bar{X}$	$\sigma_{\bar{X}} = \sigma/\sqrt{n}$	$S/\sqrt{n}$
μ_1 - μ_2	$\bar{X} - \bar{Y}$	Si $\sigma_1 \neq \sigma_2$: $\sigma_{\bar{X}-\bar{Y}} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$	$\hat{\sigma}_{\bar{X}-\bar{Y}} = \sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}$
μ_1 - μ_2	$\bar{X} - \bar{Y}$	Si $\sigma_1 = \sigma_2$: $\sigma_{\bar{X}-\bar{Y}} = \sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$	$\hat{\sigma}_{\bar{X}-\bar{Y}} = S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$ où $S = \sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$
p	$\hat{p}$ (proportion échantillonnale)	$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$	$\hat{\sigma}_{\hat{p}} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$
p_1 - p_2	$\hat{p}_1 - \hat{p}_2$	$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$	$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
p_1 - p_2	$\hat{p}_1 - \hat{p}_2$	Sous H_o : $p_1 = p_2 = p$ $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$	$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\hat{p}(1-\hat{p})}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$ où $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1+n_2}$