IMBS3 - ISCID-CO, site de Dunkerque, 2015/2016

Mathématiques appliquées à la gestion - TESTS PARAMÉTRIQUES ET NON PARAMÉTRIQUES

Fiche de Mathématiques 2 - Notions générales sur les tests.

Exercice 1 *Correction* :
La variable aléatoire X suit une loi binomiale telle que $X \rightsquigarrow B(n,p)$ avec n le nombre de lancers d'où $n = 100$ et p est la probabilité du succès "obtenir face" donc $p = \frac{1}{2}$ si la pièce n'est pas truquée. La loi de probabilité suivie par X implique $E(X) = np = 50$ et $V(X) = npq = 25$. Le risque d'erreur de première espèce α correspond à $P(\text{rejeter } H_0/H_0 \text{ vraie}) = P(X \notin [40,60])$. Par conséquent, $P(40 \leq X \leq 60) = 1 - \alpha$. En pratique, on sait qu'une loi binomiale $\mathcal{B}(n,p)$ peut être approchée par une loi normale $\mathcal{N}(np, \sqrt{npq})$ si $n \geq 20$, $np \geq 10$ et $nq \geq 10$, ce qui est le cas ici. On centre et on réduit X, on a ainsi $P(40 \leq X \leq 60) = P\left(\frac{40-50}{\sqrt{25}} \leq T \leq \frac{60-50}{\sqrt{25}}\right) = P(-2 \leq T \leq 2) = 2P(T \leq 2) - 1 = 0,954 = 1 - \alpha$ avec $P(T \leq 2) = 0,977$ (table de la loi normale centrée réduite). Finalement, le risque d'erreur de première espèce α est égal à $0,046$. On a $4,6\%$ de chance d'accepter l'hypothèse H_1 : "la pièce est truquée" alors que la pièce n'est pas truquée (H_0 vraie).

Exercice 2 *Correction* :
La variable aléatoire X suit toujours une loi binomiale $\mathcal{B}(n,p)$ mais avec des paramètres différents. Dans notre cas, n est le nombre de lancers d'où $n = 100$ et p est la probabilité du succès "obtenir face" d'où $p = 0,6$ si la pièce est truquée. La loi de probabilité suivie par X implique $E(X) = np = 60$ et $V(X) = npq = 24$. Le risque d'erreur de seconde espèce β correspond à $P(\text{accepter } H_0/H_1 \text{ vraie}) = P(X \in [40,60]) = P(40 \leq X \leq 60)$. $\mathcal{B}(n,p)$ peut une nouvelle fois être approchée par une loi normale $\mathcal{N}(np, \sqrt{npq})$ puisque $n \geq 20$, $np \geq 10$ et $nq \geq 10$. On centre et on réduit X, ainsi $\beta = P\left(\frac{40-60}{\sqrt{24}} \leq T \leq \frac{60-60}{\sqrt{24}}\right) = P(-4,08 \leq T \leq 0) = P(0) - P(-4,08) = 0,5$ avec $P(0) = 0,5$ et $P(-4,08) = 1 - 0,9999$ (table de la loi normale centrée réduite). Le risque d'erreur de seconde espèce β est $0,50$. On a 50% de chance d'accepter l'hypothèse H_0 : "la pièce n'est pas truquée" alors que la pièce est truquée (H_1 vraie).

Exercice 3 *Correction* :
Le test porte sur des hypothèses de recherche.

1. Si μ est le volume moyen des ventes associé au nouveau système de bonus, on considérera H_0 : "$\mu \leq 14$", H_1 : "$\mu > 14$".
2. Aucune preuve que le nouveau projet accroisse les ventes.
3. L'hypothèse de recherche H_1 : "$\mu > 14$" est acceptée, le nouveau projet accroît les ventes.

Exercice 4 *Correction* :
La variable aléatoire X suit une loi binomiale telle que

Valeur de p	$E(X)$	$V(X)$	$\beta = P(40 \leq X \leq 60)$	Puissance $1-\beta$
0,3	30	21		
0,4	40	24	0,5	0,5
0,6	60	24	0,5	0,5
0,7	70	21	0,0147	0,9853
0,8	80	16	0	1
0,9	90	9	0	1

Les différentes probabilités p permettent d'établir différentes lois de probabilité de X sous "H_1 est vraie" (pièce de monnaie truquée) et donc d'établir différentes valeurs du risque de seconde espèce β. Le risque α reste fixé à $0,05$. Le risque d'erreur de seconde espèce β correspond à $P(\text{accepter } H_0/H_1 \text{ est vraie}) = \beta$ d'où $P(40 \leq X \leq 60) = \beta$.
La puissance du test $1-\beta$ est d'autant plus forte que l'hypothèse H_1 est différente de H_0.

Pour avoir une chance de gagner en trichant sans que l'autre s'en doute, il faut que la pièce ne soit pas trop truquée : $p = 0,5 \pm 0,1$. Lorsque la pièce est trop truquée, la probabilité d'accepter H_0 est très inférieure à 2%.

Exercice 5 *Correction* :
Soit X_i la variable aléatoire qui à un individu i associe son choix de vote, alors $X_i \rightsquigarrow \mathcal{B}(p)$ où p désigne la fréquence d'intention de vote.

1. Sous H_0, $X_i \rightsquigarrow \mathcal{B}(0,39)$. Alors $X = \sum_{i=1}^{100} X_i \rightsquigarrow \mathcal{B}(100;0,39)$. Les conditions de normalité sont vérifiées : $n = 100 \geq 30$, $np = 100 \times 0,39 = 39 \geq 5$ et $nq = 100 \times 0,61 = 61 \geq 5$ donc $\mathcal{B}(100;0,39) \approx \mathcal{N}(100 \times 0,39; \sqrt{100 \times 0,39 \times 0,61}) = \mathcal{N}(39;4,88)$ (l'approximation résulte de la convergence en loi de la loi binomiale vers la loi normale, cas particulier du théorème central limite). Par conséquent, $F = \frac{X}{n} \rightsquigarrow \mathcal{N}(0,39; \sqrt{\frac{0,39\times0,61}{100}}) = \mathcal{N}(0,39;0,0488)$ et $Z = \frac{F-0,39}{0,0488} \rightsquigarrow \mathcal{N}(0,1)$. Alors, si f_n est la fréquence d'intention de vote obtenue à l'aide de l'échantillon, $\alpha = P(\text{rejeter } H_0/H_0 \text{ est vraie}) = P(F \geq f_n) = P(Z \geq Z_\alpha)$ où $Z_\alpha = \frac{f_n-0,39}{0,0488}$.
2. Si $\alpha = 5\% = 0,05$, $1-\alpha = 0,95$ et $Z_{0,05} = 1,645 = \frac{f_n-0,39}{0,0488} \Leftrightarrow f_n = 0,39+1,645\times0,0488 = 0,4703$. Conclusion, la région critique est $F = \frac{1}{n}\sum_{i=1}^{n} X_i \geq 0,4703$. Si dans l'échantillon, le nombre d'individus votant pour Mr X est supérieur à 47, on rejette H_0.
3. On rappelle que $\beta = P(\text{rejeter } H_1/H_1 \text{ est vraie})$. Sous H_1, $F \rightsquigarrow \mathcal{N}(0,41;0,0492)$. Ainsi, sous H_1, si $Z = \frac{F-0,41}{0,0492}$, $\beta = P(F \leq 0,4703) = P(Z \leq 1,23) = 0,8907$.
La puissance statistique se définit par l'aptitude du test (en termes de probabilité) à obtenir un résultat statistiquement significatif si l'expérience est réellement efficace. Ici, le test a 11 chances sur 100 de détecter une intention de vote de 41% (ou plus). Mr X ne peut pas réellement faire confiance à l'institut de sondage. Les valeurs p_0 et p_1 étant très proches, la probabilité pour que le test accepte H_0 alors que H_1 est vraie est importante, ceci du fait de l'éloignement de la valeur seuil de ces deux valeurs.
4. $f_n = 0,4 < 0,47$ donc on conserve H_0 au risque $\alpha = 5\%$: on a 5 chances sur 100 de rejeter H_0 - et donc donner raison à l'institut de sondage - alors que H_0 est vraie. Cependant, on a 89 chances sur 100 d'accepter H_0 alors que H_1 est vraie.
5. $n = 100$ permet actuellement de ne pas accepter H_1 au risque $\alpha = 5\%$ mais avec un très grand risque β. Il faut donc utiliser un échantillon dont la taille n permet d'augmenter la puissance du test. On cherche n tel que la région critique soit définie par $F = \frac{1}{n}\sum_{i=1}^{n} X_i < 0,4$. Alors $f_n = 0,4 = 0,39 + 1,645 \times \sqrt{\frac{0,39\times0,61}{n}} \Leftrightarrow n = 6438$. Alors, sous H_1, $F \rightsquigarrow \mathcal{N}(0,41; \sqrt{\frac{0,41\times0,59}{6438}}) = \mathcal{N}(0,41;6,13\times10^{-3})$. En posant $Z = \frac{F-0,41}{6,13\times10^{-3}}$, $\beta = P(F \leq 0,405) = P(Z \leq -0,816) = 1 - P(Z \leq 0,816) = 1 - 0,79 = 0,21$.
6. Pour le niveau $\alpha = 5\%$ fixé, calculons le nombre d'observations nécessaire pour obtenir la puissance donnée. Sous H_0, $F \rightsquigarrow \mathcal{N}(0,39; \sqrt{\frac{0,39\times0,61}{n}})$. Comme $\alpha = 0,05 = P_{H_0}(Z \geq 1,645)$ avec $Z = \frac{F-0,39}{\sqrt{\frac{0,39\times0,61}{n}}}$, on a $Z_\alpha = 1,645 = \frac{f_n-0,39}{\sqrt{\frac{0,39\times0,61}{n}}}$ soit $f_n = 1,645\sqrt{\frac{0,39\times0,61}{n}} + 0,39$.
$\beta = P(\text{rejeter } H_1/H_1 \text{ est vraie})$. Sous H_1, $F \rightsquigarrow \mathcal{N}(0,42; \sqrt{\frac{0,42\times0,58}{n}})$. Comme $\beta = 0,1 = P(\frac{F-0,42}{\sqrt{\frac{0,42\times0,58}{n}}} \leq -1,29)$, on a $-1,29 = \frac{f_n-0,42}{\sqrt{\frac{0,42\times0,58}{n}}}$ soit $f_n = -1,29\sqrt{\frac{0,42\times0,58}{n}} + 0,42$. On a par conséquent $1,645\sqrt{\frac{0,39\times0,61}{n}} + 0,39 = -1,29\sqrt{\frac{0,42\times0,58}{n}} + 0,42 \Leftrightarrow n \geq 2301$.

Exercice 6 *Correction* :

1. Le risque de première espèce est défini par $P(\text{rejeter } H_0/H_0 \text{ est vraie})$. Soit F l'estimateur de la proportion de grévistes convaincus. Alors $\alpha = P_{H_0}(F < l)$ (test unilatéral gauche) où $F \rightsquigarrow \mathcal{N}(0,53; \sqrt{\frac{0,53\times0,47}{100}}) = \mathcal{N}(0,53;0,05)$ et l désigne la région critique. Ainsi, $\alpha = P(Z < Z_\alpha)$ où $Z = \frac{F-0,53}{0,05} \rightsquigarrow \mathcal{N}(0,1)$ et $Z_\alpha = \frac{l-0,53}{0,05}$.
2. On pose $\alpha = 0,05$. On a d'après la table de la loi normale centrée réduite $P(Z < Z_\alpha) = 0,05 \Leftrightarrow Z_\alpha = -1,645$. Ainsi, $-1,645 = \frac{l-0,53}{0,05} \Leftrightarrow l = 0,4478$. La région critique est donc $\frac{1}{n}\sum_{i=1}^{n} X_i < 0,4478$.
3. Comme $p_n = 0,51 > 0,4478$, on ne peut pas rejeter l'hypothèse H_0.
4. $\beta = P(\text{rejeter } H_1/H_1 \text{ est vraie}) = P_{H_1}(F > l) = P(\frac{F-0,47}{0,05} > \frac{l-0,47}{0,05}) = P(\frac{F-0,47}{0,05} > -0,444) = 0,67$. L'erreur de seconde espèce est donc conséquente.

5. $\eta = 1 - \beta = 0,33$.

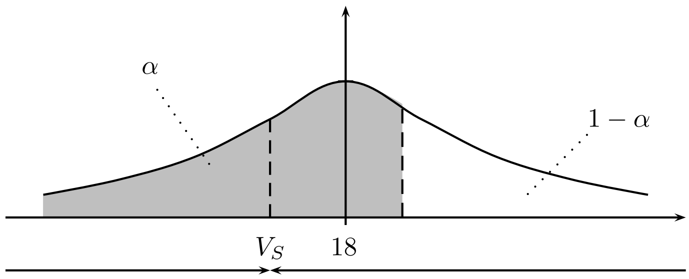
6. Pour le niveau $\alpha = 5\%$ fixé, calculons le nombre d'observations nécessaire pour obtenir la puissance donnée. Sous H_0, $F \rightsquigarrow \mathcal{N}(0,53; \sqrt{\frac{0,53\times 0,47}{n}})$. Comme $\alpha = 0,05 = P_{H_0}(Z \leq -1,645)$ avec $Z = \frac{F-0,53}{\sqrt{\frac{0,53\times 0,47}{n}}}$, on a $Z_\alpha = -1,645 = \frac{f_n - 0,53}{\sqrt{\frac{0,53\times 0,47}{n}}}$ soit $f_n = -1,645\sqrt{\frac{0,53\times 0,47}{n}} + 0,53$.
$\beta = P(\text{rejeter } H_1/H_1 \text{ est vraie})$. Sous H_1, $F \rightsquigarrow \mathcal{N}(0,47; \sqrt{\frac{0,47\times 0,53}{n}})$. Comme $\beta = 0,1 = P(\frac{F-0,47}{\sqrt{\frac{0,47\times 0,53}{n}}} \geq 1,29)$, on a $1,29 = \frac{f_n - 0,47}{\sqrt{\frac{0,47\times 0,53}{n}}}$ soit $f_n = 1,29\sqrt{\frac{0,47\times 0,53}{n}} + 0,47$. On a par conséquent $-1,645\sqrt{\frac{0,47\times 0,53}{n}} + 0,53 = 1,29\sqrt{\frac{0,47\times 0,53}{n}} + 0,47 \Leftrightarrow n \geq 597$.

Exercice 7 *Correction* :

1. On pose les hypothèses nulle et alternative sachant que $\alpha = 2\%$.

$$H_0 : \text{“Le temps d'attente n'a pas changé, } m = 18\text{” et } H_1 : \text{“Le temps d'attente a diminué, } m < 18\text{”.}$$

Sous H_0, $\overline{X}$ la variable aléatoire égale au temps d'attente sur tout échantillon de taille 100 suit une loi normale $\mathcal{N}(18, \frac{7.2}{\sqrt{100}}) = \mathcal{N}(18, 0.72)$



Zone de rejet Zone d'acceptation

Sous H_0 on cherche la valeur de a telle que $P(\overline{X} \geq 18 - a) = 0,98$, ce qui équivaut, en posant $T = \frac{\overline{X}-18}{0,72}$ où T suit la loi normale centrée réduite, à $P(T \geq \frac{-a}{0,72}) = 0,98$ soit par symétrie de la loi normale $P(T \leq \frac{a}{0,72}) = 0,98$. Donc par lecture inverse sur la table de la loi normale centrée réduite et interpolation linéaire, on obtient $\frac{a}{0,72} = 2,055 \Leftrightarrow a = 1,48$. Ainsi on peut énoncer la règle de décision :

- Si $me \geq 16,52$ alors on accepte H_0 avec un risque β inconnu.
- Si $me > 16,52$, on rejette H_0 donc on accepte H_1 avec une probabilité de risque $\alpha = 2\%$.

2. Ici la probabilité critique vaut $P(\overline{X} \geq me)$ où $me = 15,78$ sous H_0.

$$P(\overline{X} \leq 15,78) = P(T \leq -3,08) = P(T \geq 3,08) = 1 - P(T < 3,08) = 1 - 0,998962 = 0,001038$$

(par interpolation linéaire). La probabilité critique vaut ici à peu près $0,1\%$.

3. Pour conclure, on peut procéder de deux façons :

- A l'aide des tests statistiques, comme $me = 15,78 < 16,52$ on rejette H_0 et donc il y a une diminution du temps d'attente au risque 2%.
- A l'aide de la probabilité critique : comme $\alpha = 2\% > pc = 0,1\%$, on rejette H_0 et donc il y a une diminution du temps d'attente.

Exercice 8 *Correction* :

1. $P(\text{rejeter } H_0/H_0 \text{ est vraie}) = P(X_1 > 0,9/X \rightsquigarrow \mathcal{U}(0,1)) = 0,1$.

2. $P(\text{rejeter } H_0/H_1 \text{ est vraie}) = P(X_1 > 0,9/X \rightsquigarrow \mathcal{U}(0,1.5) = 0,6/1,5 = 0,4$.

Exercice 9 *Correction* :

1. On rejette H_0 si $\overline{X} < -1,645 \times 5/3 + 50 = 47,25$.
   $P(\text{rejeter } H_0 / X \rightsquigarrow \mathcal{N}(45, 25)) = P(X < 47,25) = P(Z < (47,25 - 45) \times 3/5) = P(Z < 1,35) = 0,91$.

2. $n \geq (Z_{0,05} + Z_{0,10})^2 \times 25/(45-50)^2 = (1,645 + 1,28)^2 = 8,6$. On prendra $n = 0$.