



TD4 – Corrigé

Exercice 1 :

M (l'anxiété dans l'échantillon) devrait suivre une loi normale de moyenne $\bar{M}=\mu$ et d'écart type

$$s=\sqrt{\frac{\sigma^2}{N}}=\sqrt{\frac{9}{100}}=\frac{3}{10}=0,3. \text{ On cherche à vérifier si c'est bien le cas.}$$

La statistique étudiée est l'écart entre la moyenne de la population et celle de l'échantillon

$$S=\bar{M}-\mu \text{ dont la distribution de probabilité suit une loi normale } N(0 ; \sqrt{\frac{\sigma^2}{N}}).$$

Nous pouvons établir la variable Z centrée réduite telle que $Z=\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{N}}}=\frac{\bar{X}-\mu}{s}$. On appelle e_{obs} sa

valeur absolue. L'hypothèse testée est que S est nulle. Pour ce faire, on compare e_{obs} avec une valeur seuil e_{seuil} lue dans la table qui correspond au risque choisi.

$$e_{obs}=\frac{\bar{X}-\mu}{s}=\frac{11-10}{0,3}\approx 3,33$$

Avec un risque d'erreur $\alpha = 0,05$, $e_{seuil} = 1,96$ donc $e_{obs} > e_{seuil}$ et l'hypothèse est rejetée avec un risque d'erreur de 5%. L'échantillon peut donc être considéré comme atypique.

Exercice 2 :

Avec $\pi = 1/4$ et $N=300$, $N(\pi)$ et $N(1-\pi)$ sont tous les deux supérieurs à 10. On peut donc remplacer la loi binomiale des fréquences des échantillons de 300 personnes par une loi normale de

$$\text{moyenne } \mu = \pi = 1/4 \text{ et d'écart-type } \sigma = \sqrt{\frac{\frac{1}{4} \times \frac{3}{4}}{300}} = 0,025$$

Alors au risque 5% $\varepsilon = 1,960$ et $e = \varepsilon\sigma = 0,050$

Au risque 1% $\varepsilon = 2,576$ et $e = \varepsilon\sigma = 0,0644$

Exercice 3 :

La variable aléatoire X correspondant au fait de jouer régulièrement aux jeux vidéos suit une loi binomiale de paramètres N, π . Le nombre de joueurs réguliers et de non joueurs sont supérieurs à

10 donc on peut faire une approximation par une loi normale $N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$.

Si on se donne une probabilité α de 95%, on cherche quel est l'intervalle centré sur p qui contient 95% de probabilité pour la variable aléatoire X.

$$I_{95\%} = \left[\pi - \varepsilon \sqrt{\frac{\pi(1-\pi)}{n}}; \pi + \varepsilon \sqrt{\frac{\pi(1-\pi)}{n}} \right] = \left[\pi - 1,96 \sqrt{\frac{\pi(1-\pi)}{681}}; \pi + 1,96 \sqrt{\frac{\pi(1-\pi)}{681}} \right]$$

26,7% de l'échantillon joue régulièrement. C'est une estimation ponctuelle. Pour avoir une estimation par intervalle, on considère l'échantillon comme représentatif et on utilise ses caractéristiques dans la formule.

$$I_{95\%} = \left[0,267 - 1,96 \sqrt{\frac{0,267 \times 0,733}{681}}; 0,267 + 1,96 \sqrt{\frac{0,267 \times 0,733}{681}} \right] = [0,2338; 0,3002] = [23,4\%; 30,0\%]$$

Exercice 4 :

1. Une estimation ponctuelle de la moyenne μ des diamètres des pièces P produites pendant cette journée est 4,012. C'est la moyenne au sein de l'échantillon.
2. On cherche le nombre réel positif a tel que $p(X + a \leq \mu \leq X + a) = 0,95$. X suit une loi normale $N(\mu; \frac{\sigma}{\sqrt{60}})$ donc la variable aléatoire T définie par $T = \frac{X - \mu}{\frac{\sigma}{\sqrt{60}}}$ suit la loi normale centrée réduite.

Donc on a $p(-t \leq T \leq t) = 0,95$, soit $t = 1,96$

$$p(-t \leq T \leq t) = p\left(-t \leq \frac{X - \mu}{\frac{\sigma}{\sqrt{60}}} \leq t\right) = p\left(-t \frac{\sigma}{\sqrt{60}} \leq X - \mu \leq t \frac{\sigma}{\sqrt{60}}\right) = p\left(X - t \frac{\sigma}{\sqrt{60}} \leq \mu \leq X + t \frac{\sigma}{\sqrt{60}}\right)$$

En prenant $\bar{x} = 4,012$, $\sigma = 0,084$, et $t = 1,96$ on en déduit que [3,991 ; 4,033] est un intervalle de confiance centré en \bar{x} de la moyenne μ des diamètres des pièces P produites pendant la journée considérée, avec le coefficient de confiance de 95%.

3. On considère l'affirmation suivante : " la moyenne μ est obligatoirement entre 3,991 et 4,033 ". On ne peut pas en déduire que ce qui précède est vrai, il n'y a pas de certitude.

Exercice 5 :

On cherche à déterminer si l'échantillon est représentatif d'une population de livraisons correspondant aux engagements du fournisseur. Le but est donc de savoir si un échantillon de fréquence observée p appartient à la population de référence connue de fréquence p_0 . ou à une autre population inconnue. La fréquence dans l'échantillon suit une loi normale de paramètres p et $\sqrt{\frac{p_0(1-p_0)}{N}}$, les variances étant supposées égales entre la population d'où est extrait l'échantillon et la population de référence.

Nous pouvons établir la variable Z centrée réduite telle que $Z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{N}}}$. On peut appeler e_{obs} la

valeur absolue de cette variable. L'hypothèse testée est que $p - p_0$ est nul. Pour ce faire, on compare e_{obs} avec une valeur seuil e_{seuil} lue dans la table qui correspond au risque choisi.

$$e_{obs} = \frac{|p - p_0|}{\sqrt{\frac{p_0(1-p_0)}{N}}} = \frac{|0,85 - 0,9|}{\sqrt{\frac{0,9 \times 0,1}{300}}} \approx 2,88$$

Avec un risque d'erreur $\alpha = 0,05$, $e_{seuil} = 1,96$ donc $e_{obs} > e_{seuil}$ et l'hypothèse est rejetée avec un risque d'erreur de 5%. On peut considérer que le fournisseur ne respecte pas ses engagements.

Exercice 6 :

$n > 30$, la variable aléatoire P : "proportion de mauvais tests dans un échantillon de 400 éléments" est approximativement distribuée selon la loi normale : $N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{400}}\right)$

La règle de décision est la suivante : si on note p la proportion de mauvais tests observés dans l'échantillon, si $p < 0,06$ on considère le test comme satisfaisant, sinon, il ne l'est pas.

$$\alpha = p(p > 0,06) = p\left(U > \frac{0,06 - 0,05}{\sqrt{\frac{0,05 \times 0,95}{400}}}\right) = p(U > 0,92) = 0,1788$$

Le risque de considérer le test comme non satisfaisant est d'environ 0,18.

Exercice 7 :

On veut évaluer si $p_1 = p_2$, soit si $d = p_1 - p_2$ peut être considéré comme égal à 0. On fait un test de seuil 10%.

$$\varepsilon' = \frac{|p_1 - p_2|}{\sqrt{\frac{p(1-p)}{N_1} + \frac{p(1-p)}{N_2}}} = \frac{|p_1 - p_2|}{\sqrt{\frac{N_1 p_1 + N_2 p_2}{N_1 + N_2} \left(1 - \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2}\right) + \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2} \left(1 - \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2}\right)}}$$

$$\varepsilon' = \frac{\left|\frac{29}{250} - \frac{49}{350}\right|}{\sqrt{\frac{0,13 \times 0,87}{250} + \frac{0,13 \times 0,87}{350}}} = \frac{0,024}{\sqrt{0,000776}} = 0,862$$

Avec un seuil de 10%, on doit avoir, si l'on fait l'hypothèse d'une loi normale, $p(U > \varepsilon) = 0,10$, soit $p(|U| > \varepsilon) = 0,20$

Or $0,862 < 1,282$, donc au seuil de 10%, la proportion de retour favorable chez les plus de 35 ans n'est pas significativement supérieure à celle observée chez les moins de 35 ans.

Exercice 8 :

On veut évaluer si $p_1 = p_2$, soit si $d = p_1 - p_2$ peut être considéré comme égal à 0. On fait un test de seuil 10%.

$$\varepsilon' = \frac{|p_1 - p_2|}{\sqrt{\frac{p(1-p)}{N_1} + \frac{p(1-p)}{N_2}}} = \frac{|p_1 - p_2|}{\sqrt{\frac{s_1^e}{140} + \frac{s_2^e}{110}}}$$

$$\varepsilon' = \frac{|141,6 - 160,86|}{\sqrt{\frac{1973}{140} + \frac{9085}{110}}} = 1,96$$

Avec un seuil de 5%, on doit avoir, si l'on fait l'hypothèse d'une loi normale, $p(U > \varepsilon) = 0,05$, soit $p(|U| > \varepsilon) = 0,10$

Or $1,96 > 1,645$, donc la différence observée est significative. Le capital nécessaire à la création est en moyenne différent selon le degré technologique de l'entreprise.