

TD Statistique inférentielle¹

Interro 7-8

Une société pharmaceutique a développé un test sérologique (détectant la présence d'anticorps) permettant de tester si une personne est immunisée au COVID-19. On suppose qu'on connaît la probabilité ε_1 que le test soit négatif pour une personne ayant l'immunité au COVID-19, ainsi que la probabilité ε_2 que le test soit positif pour une personne qui n'a pas d'immunité.

On note p la proportion des personnes immunisées à Villeneuve d'Ascq et on suppose que l'on observe un échantillon X_1, \dots, X_n de résultats de tests sérologiques de n résidents à Villeneuve d'Ascq, avec X_i qui vaut 1 si le test de i -ème personne est positif et 0 sinon.

- (Élémentaire) Montrez que la probabilité que le test soit positive est égale à $p + \varepsilon_2 - p\varepsilon_1 - p\varepsilon_2$. Calculez p_1 la probabilité qu'une personne testée positive ait effectivement l'immunité et p_2 la probabilité qu'une personne testée négative n'ait effectivement pas d'immunité. Application numérique : calculez les valeurs de ces probabilités si $p = 95\%$, $\varepsilon_1 = 10\%$ et $\varepsilon_2 = 1\%$. Que remarquez-vous ?
- (Élémentaire) Construire un estimateur p_n^* de p par la méthode des moments.
- En s'inspirant de EX 2 du TD 3, trouvez un intervalle de confiance $I_n = I_n(X_1, \dots, X_n)$ de niveau de risque 0,05 pour p .
- (Bonus) Refaites la question précédente sans l'approximation au dénominateur.
- (Élémentaire) Pour une réalisation x_1, \dots, x_n , on trouve que $I_n(x_1, \dots, x_n) = [0,63; 0,66]$, peut-on dire que "la proportion des villeneuvois qui sont immunisés au COVID-19 est comprise entre 63% et 66% avec une probabilité de 95%" ? Expliquez.

On dit qu'il y a une "immunité collective" si $p > p_0$, où p_0 est une constante connue (par exemple 60%). Pour tester (statistiquement) s'il y a bien une immunité collective, on considère donc dans la suite un test statistique de l'hypothèse nulle " $p = p_0$ " contre l'alternative " $p > p_0$ " basé sur la variable de décision \bar{X}_n .

- (Élémentaire) Choisissez (en justifiant votre choix) la forme de la zone de rejet parmi $]-\infty, k]$ et $[k, +\infty[$.
- En utilisant l'approximation normale de la loi Binomiale, déterminez k (on fixe le risque de première espèce égale à $\alpha = 0,01$).
- Exprimez, à l'aide de la fonction de répartition de la loi normale centrée réduite Φ , le risque de seconde espèce (en fonction de $p \in]p_0, 1[$).
- On suppose que $p_0 = 60\%$, $\varepsilon_1 = 10\%$ $\varepsilon_2 = 1\%$ et que sur 1000 personnes testées, 585 ont été positives. Si vous étiez maire de Villeneuve d'Ascq, annuleriez-vous les mesures sanitaires ? Expliquez.

1. Mohamed-slim.kammoun@univ-lille.fr