Réponses pour les premiers exercices du TD Intervalles de confiance

Je n'ai pas fait toutes les applications numériques.

1. (a) $\bar{x} = 72.37$
   $s = \sqrt{\frac{\sum_{i=1}^{10}(X_i - 72.37)^2}{9}} = 0.117$
   (b) L'intervalle est de la forme $[\bar{x} - z_\alpha \frac{s}{\sqrt{10}}, \bar{x} + z_\alpha \frac{s}{\sqrt{10}}]$ où z_α est tel que $P(|Z| > z_\alpha) = \alpha$, si Z suit une loi de Student à 9 degrés de liberté. On a donc $z_{0.1} = 1.83$ et $z_{0.05} = 2.26$.
   Soit $I_{0.1} = [72.37 - 1.83 * 0.117/\sqrt{10}, 72.37 + 1.83 * 0.117/\sqrt{10}]$ et $I_{0.1} = [72.37 - 2.26 * 0.117/\sqrt{10}, 72.37 + 2.26 * 0.117/\sqrt{10}]$.
   (c) Autre estimation de l'espérance $\bar{x}_3 = \frac{\bar{x}+\bar{x}_2}{2} = 72.35$. Pour lécart-type, on a besoin de $\sum_{i=1}^{20}(X_i - \bar{x}_3)^2$. On utilise $\sum_{i=1}^{20}(X_i - \bar{x}^3)^2 = \sum_{i=1}^{20} X_i^2 - 20\bar{x}_3^2$, et $\sum_{i=1}^{10} X_i^2 = 9 * 0.117^2 + 10\bar{x}^2 = 9 * 0.117^2 + 10 * 72.37^2$ et $\sum_{i=11}^{20} X_i^2 = 9 * 0.07^2 + 10\bar{x}^2 = 9 * 0.07^2 + 10 * 72.33^2$.
   Finalement $\sum_{i=1}^{20}(X_i - \bar{x}_3)^2 = 0.175$ soit un nouvel écart-type de 0.096, soit à peu près 0.1.
   (d) L'intervalle de confiance est $[\bar{x} - \sigma\frac{t_\alpha}{\sqrt{n}}, \bar{x} + \sigma\frac{t_\alpha}{\sqrt{n}}]$ où t_α est tel que $P(|Y| > t_\alpha) = \alpha$ si Y suit une loi normale centrée réduite.
   Soit $I_{0.05} = [72.35 - 0.08 * 1.96/\sqrt{20}, 72.35 + 0.08 * 1.96/\sqrt{20}]$.
   (e) L'intervalle proposé n'est pas centré autour de $\bar{x}_3$, c'esty une erreur (qui vient d'une version antérieure de l'exercice). On va y répondre avec $[72.3, 72.4]$. On aurait alors $z_\alpha * 0.08/\sqrt{20} = 72.35 - 72.3 = 0.05$ soit $z_\alpha = 2.8$, d'où $\alpha = 0.005$.

2. Le nombre de moutons guéris est une binomiale de paramètre $(100, p)$ où p est la probabilité de guérison. Elle a la même loi qu'une somme de 100 bernoulli indépendantes de paramètre p. La probabilité de guérison d'un mouton (i.e. l'espérance de la Bernoulli) est donc estimée par 63/100. D'autre part $\frac{\text{Nombre moutons guéris} - 100*p}{\sqrt{100*p*(1-p)}}$ suit à peu près une loi normale centrée réduite.

   Deux possibilités : version facile : on approche l'écart-type $\sqrt{p(1-p)}$ par $0.63 * (1 - 0.63)$. On a donc pour intervalle de confiance pour p $[0.63 - t_\alpha * 0.63 * (1 - 0.63)/10, 0.63 + t_\alpha * 0.63 * (1 - 0.63)/10]$ où t_α est lu dans la table de la loi normale et vaut 1.96. Soit $I = [0.58, 0.68]$.

   Version plus correcte : on a $P(|\text{Nbre moutons guéris} - 100p| \geq 1.96 * 10\sqrt{p(1-p)}) = 0.05$. Or $|N - 100p| \geq 19.6 * \sqrt{p(1-p)}$ si et seulement si $N^2 + 10000p^2 - 200pN \geq 19.6^2 * (p - p^2)$, c'est à-dire si et seulement si

   $N^2 + (10000 + 19.6^2)p^2 - p(200N + 19.6^2) \geq 0$, donc si et seulement si p est à l'extérieur des racines de ce polynôme. Ici, $N = 63$ et on obtient donc comme intervalle de confiance $[0.53, 0.72]$: il est plus gros que le précédent (et plus correct...)

Si on veut une longueur de 0.01, avec la méthode “facile”, on obtient qu'on doit avoir $2*1.96*\sqrt{p(1-p)}/\sqrt{n} = 0.01$. Donc $n = 4*1.96^2*p(1-p)/0.0001$. En gardant $p = 0.63$, on obtient $n = 35820$ moutons... Au pire, $p(1-p)$ est toujours inférieur à 0.25 (car p est compris entre 0 et 1) et on trouve $n = 38416$, Ce qui est beaucoup...

3. Ici $n = 1$, et $I = [2.7 - 1.96*\sqrt{0.01}, 2.7 + 1.96*\sqrt{0.01}] = [2.5, 2.9]$. L'intervalle contient 3 pour $2.7 + t_\alpha\sqrt{0.01} \geq 3$ soit $t_\alpha \geq 3$, soit $\alpha \leq 0.002$: c'est petit, donc on est raisonnablement sûr que $m \leq 3$.

4. Ici, on utiliser Student, avec 59 degrés de liberté. Ce n'est pas dans la table, mais on trouve un $z_{0.05}$ de l'ordre de 2, soit $I = [199-2*\frac{\sqrt{1252}}{\sqrt{60}}, 199-2*\frac{\sqrt{1252}}{\sqrt{60}}] = [197.82, 200, 18]$. Si on garde la loi normale, 2 est remplacé par 1.96, ce qui ne change pas grand chose.

5. a) Une estimation de p est $24/200 = 0.12$. On peut prendre comme intervalle de confiance (méthode simple) $I_{0.05} = [0.12 - 1.96*\sqrt{\frac{0.12*0.88}{200}}, 0.12 - 1.96*\sqrt{\frac{0.12*0.88}{200}} = [0.08, 0.165]$.

   $I_{0.01} = [0.12 - 2.576*\sqrt{\frac{0.12*0.88}{200}}, 0.12 - 2.576*\sqrt{\frac{0.12*0.88}{200}} = [0.06, 0.18]$.

   b) On devrait avoir $2*1.96*\sqrt{\frac{p(1-p)}{n}} = 0.04$. Si on pense que 0.12 est un bon ordre de grandeur, on peut le garder et obtenir $n = 4*1.96^2*0.12*0.88/0.04^2 = 1014$.

6. a)Les intentions de vote pour C sont d'environ 212/400 : le fait qu'une personne prise au hasard vote ou non pour C est considéré comme suivant une loi de Bernoulli de paramètre p inconnu, estimé par la moyenne arithmétique 212/400. Un intervalle de confiance pour p est donc (version simple)

   $[212/400 - 1.96*\sqrt{\frac{\frac{212}{400}(1-\frac{212}{400})}{400}}, 212/400 + 1.96*\sqrt{\frac{\frac{212}{400}(1-\frac{212}{400})}{400}}] = [0.48, 0.58]$.

   b). Pour cela, il faut avoir $1.96*\sqrt{0.53*0.47}/\sqrt{(n)} \leq 0.03$ soit $n \geq 1.96^2*0.53*047/0.03^2$, et finalement $n \geq 1064$.