On note $D$ la variable aléatoire telle que :

$$D = \overline{X_1} - \overline{X_2}.$$

L'hypothèse nulle est $H_0 : \mu_1 = \mu_2$.

L'hypothèse alternative est $H_1 : \mu_1 \neq \mu_2$.

Le seuil de signification est fixé à 5%.

On admet que sous l'hypothèse nulle $H_0$ la variable aléatoire $D$ suit la loi normale :

$$\mathcal{N}\left(0\ ;\ \frac{25^2 + 20^2}{200}\right).$$

**1.** Sous l'hypothèse nulle $H_0$ déterminer le nombre réel positif $h$ tel que :

$$P(-h \leqslant D \leqslant h) = 0{,}95.$$

**2.** Énoncer la règle de décision du test.

**3.** On prélève un échantillon aléatoire de 200 fiches dans chacun des fichiers. La moyenne observée sur l'échantillon du fichier 1 est $\overline{x_1} = 133$. Celle observée sur l'échantillon du fichier 2 est $\overline{x_1} = 130$. Peut-on, au seuil de signification de 5 %, accepter l'hypothèse $H_0$ ?

**45** Pour un groupe de 300 personnes bien portantes, on a dosé le cholestérol et obtenu les résultats résumés dans le tableau suivant :

| Classe de $x_i$ | Effectif $n_i$ |
|---|---|
| [80 ; 120[ | 7 |
| [120 ; 160[ | 54 |
| [160 ; 200[ | 110 |
| [200 ; 240[ | 72 |
| [240 ; 280[ | 46 |
| [280 ; 320[ | 8 |
| [320 ; 360[ | 3 |

où $x_i$ désigne le taux de cholestérol exprimé en cg/L.

**1.** Calculer des valeurs approchées de la moyenne $\overline{x_1}$ et de l'écart type $s_1$ de l'échantillon.

**2.** Donner une estimation ponctuelle $\mu_1$ de la moyenne et $\sigma_1$ de l'écart type du taux de cholestérol chez les gens bien portants de la région considérée.

Les résultats seront donnés à $10^{-1}$ près.

**3.** Dans une autre région, un hôpital a obtenu pour un échantillon de 250 personnes une moyenne $\overline{x_2} = 191{,}2$ et un écart type $s_2 = 45{,}2$.

On suppose que toutes les analyses effectuées sont indépendantes.

On désigne par $\overline{X_1}$ la variable aléatoire qui, à chaque échantillon de 300 personnes de la première région, associe la moyenne de l'échantillon et par $\overline{X_2}$ la variable aléatoire qui, à chaque échantillon de 250 personnes de la deuxième région, associe la moyenne de l'échantillon.

On suppose que les variables aléatoires $\overline{X_1}$, $\overline{X_2}$, $D = \overline{X_1} - \overline{X_2}$ suivent les lois normales de moyennes respectives $\mu_1$, $\mu_2$, $\mu_1 - \mu_2$ inconnues et on estime l'écart type de $D$ par :

$$\sqrt{\frac{{\sigma_1}^2}{300} + \frac{{\sigma_2}^2}{250}}$$

où $\sigma_1$ et $\sigma_2$ sont les écart types estimés à partir des échantillons précédents.

On désire construire un test permettant de déterminer si il y a une différence significative entre les moyennes des deux populations au seuil de 5 %.

**a)** L'hypothèse $H_0$ est donnée par $\mu_1 = \mu_2$ ; énoncer l'hypothèse alternative $H_1$.

**b)** Déterminer l'intervalle [– a ; a] tel que, sous l'hypothèse $H_0$, $P(-a \leqslant D \leqslant a) = 0{,}95$.

**c)** Énoncer la règle de décision du test.

**d)** Utiliser ce test avec les deux échantillons de l'énoncé et conclure.

## Test de comparaison – Comparaison de deux proportions

Fiche l'Essentiel

**46 R** Les nouveaux modèles de téléviseurs que va fabriquer une usine sont de deux types : modèle (1) et modèle (2). Une enquête préalable à la fabrication, réalisée auprès de 400 ménages de la population $S$ des ménages des « quartiers sud » de la ville $V$, indique qu'entre les deux modèles de téléviseurs, 63 % préfèrent le modèle (1). La même enquête, réalisée auprès de 500 ménages de la population $N$ des ménages des « quartiers nord » de la ville, indique que 67 % préfèrent le modèle (1). On note $F_S$ la variable aléatoire qui, à tout échantillon de