Contrôle continu de novembre 2010, exercice 1

1)

L'énoncé nous informe sur la proportion de femmes chez les sapeurs pompiers volontaires.

On peut définir sur l'univers des sapeurs pompiers volontaires une variable aléatoire X qui vaut 0 si un pompier tiré au hasard est un homme et 1 si c'est une femme.

X suit une loi de Bernouilli. Comme13 % des pompiers sont des femmes, on peut considérer qu'on a 13 % de chances qu'un sapeur pompier volontaire tiré au hasard soit une femme.

On connaît donc le paramètre de la loi de X : $\mathcal{L}(X) = \mathcal{B}(1\ ;\ 0{,}13)$

2)

La proportion de femmes est faible mais elle a doublé entre 1999 et 2010.
L'activité de sapeur-pompier volontaire est masculine à l'origine, mais comme pour la plupart des métiers masculins on assiste à un mouvement de féminisation.
Ce mouvement est certainement lié à la médiatisation, par divers canaux, sous diverses formes et à divers degrés, du problème de la parité entre hommes et femmes, qui s'est accélérée ces quinze dernières années.
On peut y voir le succès lent, progressif et inabouti de l'action idéologique continue de groupes féministes depuis un siècle au moins, lui-même conséquence de différents processus, au premier plan desquels on trouve certainement ceux qui résultent, d'une part, des politiques de massification scolaire et d'autre part, des politiques de contrôle de la procréation.

3)

Numérotons de 1 à 10 les SPV interrogés, et appelons X_1 … X_{10} les variables aléatoires de Bernouilli qui indiquent s'il s'agit de femmes.

La variable aléatoire qui donne le nombre de femmes dans l'échantillon est $S_{10} = \sum_{i=1}^{10} X_i$

Rigoureusement, S_{10} suit une loi hypergéométrique, puisqu'il y a tirage sans remise :
En effet, le nombre de femmes est de 5 996 × 0,13 = 779,
donc quand j'aurai interrogé le premier SPV :

- Si c'est une femme, il n'y aura plus que 778 chances sur 5995 que ce soit une femme.
- Si c'est un homme, il y aura 779 chances sur 5995 que ce soit une femme.

Les variables X_i ne sont pas indépendantes.

$\mathcal{L}(S_{10}) = \mathcal{H}(5996\ ;\ 10\ ;\ 0{,}13)$

Cette variable peut avoir 11 valeurs différentes : de 0 à 10 femmes dans l'échantillon.

Pour k compris entre 0 et 10, la formule qui nous donne la probabilité que le nombre de femmes de l'échantillon soit égal à k est la suivante :

$$P(S_{10} = k) = \frac{C_{779}^{k} \times C_{5217}^{10-k}}{C_{5996}^{n}}$$

Les formules du cours nous donnent l'espérance et l'écart type :

$E[S_{10}] = np = 10 \times 0,13 = 1,3$

$$\sigma[S_{10}] = \sqrt{npq \times \frac{N-n}{N-1}} = \sqrt{10 \times 0,13 \times 0,87 \times \frac{5986}{5995}} \approx 1,06$$

La valeur la plus probable est celle qui est la plus proche de l'espérance, à savoir une femme parmi les 10 SPV de l'échantillon.

Calcul de $P(S_{10} < 3)$

$P(S_{10} < 3) = P(S_{10} = 0 \text{ ou } S_{10} = 1 \text{ ou } S_{10} = 2)$

$= P(S_{10} = 0) + P(S_{10} = 1) + P(S_{10} = 2)$ car les trois événements sont indépendants

Ici, on peut utiliser directement la formule pour la loi hypergéométrique. Mais on obtient une valeur très fiable en approchant la loi hypergéométrique par une loi binomiale, comme ci-dessous :

Le cours nous dit que si le taux de sondage est assez grand, si n/N <1/10, alors on peut considérer que les variables X_i sont indépendantes et ont le même paramètre que X_1.

Or ici, $n/N = 10 / 5996 \approx 0,0017 < 1/10$

Donc on peut considérer que les variables X_i sont indépendantes et de même paramètre :

$\mathcal{L}(X_1) = \dots = \mathcal{L}(X_{10}) = \mathcal{B}(1 ; 0,13)$

Dans ces conditions, le cours nous dit l'on peut approcher la loi hypergéométrique par une loi binomiale : $\mathcal{L}(S_{10}) = \mathcal{B}(10 ; 0,13)$

$P(S_{10} < 3) = P(S_{10} = 0 \text{ ou } S_{10} = 1 \text{ ou } S_{10} = 2)$

$= P(S_{10} = 0) + P(S_{10} = 1) + P(S_{10} = 2)$ car les trois événements sont indépendants

$$= C_{10}^{0} 0,13^{0} \times 0,87^{10-0} + C_{10}^{1} 0,13^{1} \times 0,87^{10-1} + C_{10}^{2} 0,13^{2} \times 0,87^{10-2}$$

$$= 0,87^{10} + 10 \times 0,13 \times 0,87^{9} + \frac{10 \times 9}{2} \times 0,13^{2} \times 0,87^{8}$$

$\approx 0{,}87$

Il y a donc 87 % de chances pour qu'il y ait (strictement) moins de trois femmes dans un tel échantillon.

4)

$S_{400} = \sum_{i=1}^{400} X_i$ donne le nombre de femmes dans un échantillon de 400 SPV.

Rigoureusement, S_{400} suit une loi hypergéométrique, puisqu'il y a tirage sans remise :
En effet, le nombre de femmes est de 5 996 × 0,13 = 779,
donc quand j'aurai interrogé le premier SPV :

- Si c'est une femme, il n'y aura plus que 778 chances sur 5995 que ce soit une femme.
- Si c'est un homme, il y aura 779 chances sur 5995 que ce soit une femme.

Les variables X_i ne sont pas indépendantes.

$\mathcal{L}(S_{400}) = \mathcal{H}(5996 ; 400 ; 0{,}13)$

Les formules du cours nous donnent l'espérance et l'écart type :

$E[S_{400}] = np = 400 \times 0{,}13 = 52$ femmes

$\sigma[S_{400}] = \sqrt{npq \times \frac{N-n}{N-1}} = \sqrt{400 \times 0{,}13 \times 0{,}87 \times \frac{5596}{5995}} \approx 6{,}7$ femmes

Or, on a $n/N = 400 / 5996 \approx 0{,}07 < 1/10$: on donc peut considérer que les variables X_i sont indépendantes et ont le même paramètre que X_1.

$\mathcal{L}(X_1) = \ldots = \mathcal{L}(X_{400}) = \mathcal{B}(1 ; 0{,}13)$

Dans ces conditions, le cours nous dit l'on peut approcher la loi hypergéométrique
par une loi binomiale : $\mathcal{L}(S_{400}) = \mathcal{B}(400 ; 0{,}13)$

Mais par ailleurs, on a ici $np = 400 \times 0{,}13 = 52 > 10$ et $nq = 400 \times 0{,}87 > 10$,

et une application du théorème central limite nous dit qu'on peut approcher la loi binomiale par une loi normale :

$\mathcal{L}(S_{400}) \approx \mathcal{N}(np ; \sqrt{npq}) \approx \mathcal{N}(52 ; 6{,}7)$

Remarque : cette propriété du cours n'est pas dans le formulaire, mais doit être sue.

5)

Pour n=2099, n/N = 2099/5996 ≈ 0,35 >110

On n'a plus n/N < 1/10, donc on ne peut plus faire comme s'il y avait tirage avec remise, les variables $X_1 \dots X_{2099}$ ne sont plus indépendantes.

Donc, d'une part, on ne plus approcher la loi hypergéométrique par une loi binomiale et, d'autre part, on ne peut plus appliquer le théorème central limite.

Remarque : heureusement, les ordinateurs actuels permettent de faire des calculs sur des grands nombres, du type de $\dfrac{C_{N_1}^{k} \times C_{N-N_1}^{n-k}}{C_N^n}$ *avec* $N=5996$, $N_1=779$ *et* $n=2099$.

Contrôle continu de novembre 2010, exercice 2

1) Univers : les jours d'une (longue) période donnée.
Epreuve : choisir un jour au hasard.
Résultat : le nombre d'interventions pour le feu ou pour des explosions le jour choisi.

Soit X la variable aléatoire qui à un jour associe le nombre d'interventions pour le feu ou pour des explosions ce jour-là.
D'après l'énoncé, $\mathcal{L}(X) = \mathcal{N}(19 ; 3)$

Soit X^* la loi centrée réduite de X :

$X^* = \frac{X-19}{3}$ $\quad\quad \mathcal{L}(X^*) = \mathcal{N}(0 ; 1)$

Probabilité pour que le nombre d'interventions soit inférieur à 22

$P(X < 22) = P(X^* < \frac{22-19}{3}) = P(X^* < 1) \approx 0{,}8413$ d'après la table

Il y a donc 84 % de chances qu'il y ait moins de 22 interventions un jour au hasard.

Probabilité pour que le nombre d'interventions soit inférieur à 18

$P(X < 18) = P(X^* < \frac{18-19}{3}) = P(X^* < -1/3) = P(X^* > 1/3)$

$= 1 - P(X^* < 1/3) \approx 1 - P(X^* < 0{,}33) \approx 1 - 0{,}6293$ d'après la table

$\approx 0{,}3707$

Il y a donc 37 % de chances qu'il y ait moins de 18 interventions un jour au hasard.

Probabilité pour que le nombre d'interventions soit supérieur à 15

$P(X > 15) = P(X^* > \frac{15-19}{3}) = P(X^* > -4/3) = P(X^* < 4/3)$

$\approx P(X^* < 1{,}33) \approx 0{,}9082$ d'après la table

Il y a donc 91 % de chances qu'il y ait plus de 15 interventions un jour au hasard.

Probabilité pour que le nombre d'interventions soit compris entre 17 et 23

$P(17 < X < 23) = P(\frac{17-19}{3} < X^* < \frac{23-19}{3}) = P(-2/3 < X^* < 4/3) = P(X^* < 4/3) - P(X^* < -2/3)$

$= P(X^* < 4/3) - P(X^* > 2/3) = P(X^* < 4/3) - (1 - P(X^* < 2/3))$

$= P(X^* < 4/3) - 1 + P(X^* < 2/3) \approx P(X^* < 1{,}33) - 1 + P(X^* < 0{,}67)$

$\approx 0{,}9082 - 1 + 0{,}7486 \approx 0{,}6568$

Il y a donc 66 % de chances pour qu'un jour au hasard, il y ait entre 17 et 23 interventions.

Recherche d'un intervalle centré sur l'espérance dans lequel on a 95,44 % de chances de trouver, pour un jour au hasard, le nombre d'interventions pour le feu ou pour des explosions.

Si X suit une loi normale,
on sait que dans 95,44 % des cas, X est compris entre $\mu - 2\sigma$ et $\mu + 2\sigma$
(intervalle centré autour de la moyenne, d'amplitude 4 écarts types) :

$P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 95{,}44\ \%$

Ici, $\mu = 19$ et $\sigma = 3$:

$P(19 - 2\times 3 < X < 19 + 2\times 3) \approx 95{,}44\ \%$

C'est-à-dire : $P(13 < X < 25) \approx 95{,}44\ \%$

L'intervalle recherché est le suivant : [13 ; 25]. *($P(X \in [13 ; 25]) \approx 95{,}44\ \%$)*

<u>Remarque</u> : le cours nous donne une méthode similaire pour calculer l'intervalle centré qui renferme 68,26 % des valeurs de X : c'est celui qui a pour amplitude 2σ : $[\mu - \sigma ; \mu + \sigma]$.

2) REMARQUE : PAS CE TYPE DE QUESTION A L'EXAMEN DE MI-SEMESTRE.

On numérote de 1 à 7 les jours de la semaine considérée, et on note $X_1 \dots X_7$ les sept variables aléatoires qui donnent le nombre d'interventions pour le feu ou pour des explosions, pour chacun des sept jours de la semaine.

Ces 7 variables suivent toutes une même loi :

$\mathcal{L}(X_1) = \dots = \mathcal{L}(X_7) = \mathcal{N}(19\,;\,3)$

On admet qu'elles sont toutes les sept indépendantes.

On cherche le nombre total d'interventions dans la semaine, soit $S_7 = \sum_{i=1}^{7} X_i$

Le cours nous dit que sous ces conditions, $\mathcal{L}(S_7) = \mathcal{N}(n\mu\,;\,\sigma\sqrt{n}) = \mathcal{N}(7\times 19\,;\,3\sqrt{7})$

(Formulaire : propriété 4 au milieu de la page 2.)

$\mathcal{L}(S_7) \approx \mathcal{N}(133\,;\,7{,}94)$

Pour le calcul de probabilités, on passe par la loi centrée réduite :

$S_7^* = \dfrac{S_7 - 133}{3\sqrt{7}}$ $\qquad [\,\mathcal{L}(S_7^*) = \mathcal{N}(0\,;\,1)\,]$

$P(125 < S_7 < 141) = P\left(\dfrac{125-133}{3\sqrt{7}} < S_7^* < \dfrac{141-133}{3\sqrt{7}}\right)$

$\approx P(-1 < S_7^* < 1) \approx 68{,}26\ \%$ (valeur classique)

Il y a donc 68 % de chances qu'il y ait entre 125 et 141 interventions en une semaine.