**Exercice 1**

## Partie A

1. On reconnaît un schéma de Bernoulli dont le succès est la capture d'un poisson marqué. La probabilité du succès à chaque capture vaut $p = \dfrac{r}{N}$. Le nombre $Y_n$ de poissons marqués suit donc une loi binomiale : $\boxed{Y_n \hookrightarrow \mathscr{B}(n\,,\,p)}$

2. • $E\left(\dfrac{1}{n\,r}\,Y_n\right) = \dfrac{1}{n\,r}\,E\,(Y_n) = \dfrac{1}{n\,r}\,n\,p = \dfrac{1}{n\,r}\,n\,\dfrac{r}{N} = \dfrac{1}{N}$

   Donc $\dfrac{1}{n\,r}\,Y_n$ est un estimateur sans biais de $\dfrac{1}{N}$.

   • Pour qu'une suite d'estimateurs sans biais soit convergente, il suffit que la suite de ses variances converge vers zéro. On a :

   $V\,(Y_n) = np(1-p) = n\,\dfrac{r}{N}\left(1-\dfrac{r}{N}\right) = \dfrac{n\,r\,(N-r)}{N^2}$ d'où

$$V\left(\frac{1}{n\,r}Y_n\right) = \left(\frac{1}{n\,r}\right)^2 V\,(Y_n) = \frac{1}{n^2\,r^2}\,\frac{n\,r\,(N-r)}{N^2} = \frac{(N-r)}{n\,r\,N^2} \underset{(n\to+\infty)}{\longrightarrow} 0$$

   Donc $\dfrac{1}{n\,r}\,Y_n$ est un estimateur convergent de $\dfrac{1}{N}$.

3. Comme $P(Y_n=0)\neq 0$, $\dfrac{n\,r}{Y_n}$ aurait une probabilité non nulle de ne pas être défini : c'est pourquoi on ne peut pas le choisir comme estimateur de $N$.

## Partie B

1. (a) Soit $i$ un entier fixé de $[\![2\,,\,n]\!]$.

   • $X_i$ est le rang d'apparition du $i$-ème poisson bagué (on dit aussi que $X_i$ est le temps d'attente du $i$-ème succès dans le schéma de Bernoulli, $X_i$ suit une loi de Pascal de paramètres $i$ et $p$). Par conséquent $D_i$ représente le nombre de poissons supplémentaires, à partir du $(i-1)$-ème poisson bagué obtenu, qu'il faut capturer pour obtenir un $i$-ème poisson bagué (qui a déjà pu être pêché).

   Donc $D_i$ est le nombre de pêches nécessaires pour obtenir un poisson marqué de plus dans une suite infinie de pêches **indépendantes** (car avec remise, en supposant que les poissons ne se souviennent pas qu'il ne faut pas mordre à l'hameçon) ayant toutes une probabilité $p = \dfrac{r}{N}$ de donner un poisson marqué. Donc $D_i \hookrightarrow \mathscr{G}\left(\dfrac{r}{N}\right)$.

   • $E\,(D_i) = \dfrac{1}{p} = \dfrac{N}{r}$ et $V\,(D_i) = \dfrac{1-p}{p^2} = \dfrac{1-\frac{r}{N}}{\left(\frac{r}{N}\right)^2} = \dfrac{N\,(N-r)}{r^2}$

   (b) Comme $D_1 + D_2 + \cdots + D_n = X_n$, on obtient par linéarité de l'espérance :

   $E\,(X_n) = E\left(\displaystyle\sum_{i=1}^{n} D_i\right) = \displaystyle\sum_{i=1}^{n} E(D_i) = n\,E(D_1) = \dfrac{n}{p}$ d'où $\boxed{E(X_n) = \dfrac{n\,N}{r}}$

   De plus $D_1\,,\ D_2\,,\ \ldots\,,\ D_n$ sont supposées mutuellement indépendantes, donc

$$V\,(X_n) = V\left(\sum_{i=1}^{n} D_i\right) = \sum_{i=1}^{n} V(D_i) = n\,V(D_1) = \frac{n(1-p)}{p^2}$$

   D'où $\boxed{V\,(X_n) = \dfrac{n\,N\,(N-r)}{r^2}}$

   (c) On pose $Z_n = \dfrac{r}{n}\,X_n$.

   • $E\,(Z_n) = E\left(\dfrac{r}{n}\,X_n\right) = \dfrac{r}{n}E\,(X_n) = \dfrac{r}{n}\dfrac{n\,N}{r} = N$

   Donc $Z_n$ est un estimateur sans biais de $N$.

   • $V\,(Z_n) = V\left(\dfrac{r}{n}\,X_n\right) = \dfrac{r^2}{n^2}V\,(X_n) = \dfrac{r^2}{n^2}\dfrac{n\,N\,(N-r)}{r^2} = \dfrac{N\,(N-r)}{n}$

   On a alors $\displaystyle\lim_{n\longrightarrow+\infty} V(Z_n) = 0$. Donc $Z_n$ est un estimateur convergent de $N$.

2. $(D_i)_{i\in\mathbb{N}^*}$ est une suite de variables aléatoires, mutuellement indépendantes et suivant toutes la même loi d'espérance $\dfrac{1}{p}$ et de variance $\dfrac{1-p}{p^2}$.

   De plus $\forall\, n\in\mathbb{N}^*,\ X_n = \displaystyle\sum_{i=1}^{n} D_i$. En posant $X_n^* = \dfrac{X_n - E(X_n)}{\sqrt{V(X_n)}}$,

   le théorème de la limite centrée permet d'affirmer que la suite $(X_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire de loi normale centrée réduite $\mathcal{N}(0\,,\,1)$.

   Ce qui nous permet de dire, que pour $n$ assez grand ($n>50$), $X_n$ suit approximativement la loi normale $\mathcal{N}\left(\dfrac{n\,N}{r}\,,\,\dfrac{\sqrt{n\,N\,(N-r)}}{r}\right)$

3. Ici $r=200$, $n=50$ et $X_{50}(\omega)=450$. $\sigma^2 = V(Z_{50}) = \dfrac{N(N-200)}{50}$

   d'où $\boxed{\sigma = \dfrac{\sqrt{N(N-200)}}{5\sqrt{2}}}$ Donc $Z_{50}(\omega) = \dfrac{200}{50}X_{50}(\omega) = 1800$ constitue une estimation ponctuelle de $N$.

**Exercice 2**

## Partie A

| $x_i$ | 58,8 | 58,9 | 59,0 | 59,1 | 59,2 | 59,3 | 59,4 | 59,5 |
|---|---|---|---|---|---|---|---|---|
| $n_i$ | 1 | 3 | 3 | 5 | 5 | 4 | 3 | 1 |

1. On pose $n = 25$ et $p = 8$ ($p$ est le nombre de valeurs observées).

   On rappelle que $\overline{x_e} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{p} n_i\, x_i$ et que

   $$v_e = \frac{1}{n}\sum_{i=1}^{p} n_i(x_i - \overline{x_e})^2 = \left(\frac{1}{n}\sum_{i=1}^{p} n_i\, x_i^2\right) - \overline{x_e}^2$$

   À l'aide de la calculatrice, $\overline{x_e} = \dfrac{1}{25}(1\times 58,8 + 3\times 58,9 + \cdots + 1\times 59,5) = 59,156$

   et $v_e = \dfrac{1}{25}(1\times 58,8^2 + 3\times 58,9^2 + \cdots + 1\times 59,5^2) - 59,156^2 \approx 0,0321$

2. Estimation ponctuelle de l'écart-type de l'ensemble de la production :

   $$\widehat{\sigma} = \sqrt{\frac{n}{n-1}\, v_e} \approx 0,1828$$

## Partie B

Dans cette partie, $\alpha = 10\%$ et $n = 25$. La moyenne étant inconnue, elle est estimée à partir des données. On utilise la variable aléatoire $Z = \dfrac{(n-1)S^2}{\sigma^2} = \displaystyle\sum_{k=1}^{n}\left(\frac{X_k - \overline{X}}{\sigma}\right)^2$ qui suit la loi du $\chi^2_{n-1}$ à $n-1$ degrés de liberté.

On cherche des réels positifs $c$ et $d$ tels que $P\,(c \leqslant Z \leqslant d) = 1 - \alpha$.

Contrairement aux lois normale et de Student, la densité de la loi du $\chi^2$ n'est pas une fonction paire. Mais on a l'habitude de répartir le risque de façon symétrique en prenant :

$$P\,(Z \leqslant c) = \frac{\alpha}{2} \quad \text{et} \quad P\,(Z \leqslant d) = 1 - \frac{\alpha}{2}$$

La table des lois du $\chi^2$ donne : $c = 13,85$ et $d = 36,42$

Donc $0,90 = P\left(c \leqslant \dfrac{(n-1)S^2}{\sigma^2} \leqslant d\right) = P\left(\dfrac{(n-1)\,S^2}{d} \leqslant \sigma^2 \leqslant \dfrac{(n-1)\,S^2}{c}\right)$

On en déduit un intervalle de confiance aléatoire à 90% pour $\sigma^2$ :

$$I_{90\%} = \left[\frac{(n-1)\,S^2}{d}\,,\ \frac{(n-1)\,S^2}{c}\right]$$

puis un intervalle de confiance observé à 90% pour $\sigma^2$ :

$$\boxed{I_{90\%\,,\,obs} = \left[\frac{(n-1)\,\widehat{\sigma}^2}{d}\ ;\ \frac{(n-1)\,\widehat{\sigma}^2}{c}\right] = \left[\frac{n\,v_e}{d}\ ;\ \frac{n\,v_e}{c}\right] \simeq [0.0220\ ;\ 0.0579]}$$

## Partie C

Dans cette partie $\alpha = 5\,\%$ et $n = 25$.

1. Posons $\mu_0 = 59$. On va tester l'hypothèse nulle $\mathrm{H}_0\ :\ \mu = \mu_0$ contre l'hypothèse alternative $\mathrm{H}_1\ :\ \mu \neq \mu_0$ de manière à effectuer un test bilatéral.

2. 
   - *Choix de la variable de test.*
     La variance $\sigma^2$ étant inconnue, on choisit comme variable de test
     $Z = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$ avec $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{k=1}^{n}(X_k - \overline{X})^2$.
     On sait que $\overline{X} \hookrightarrow \mathcal{N}\left(\mu_0\,,\ \dfrac{\sigma}{\sqrt{n}}\right)$ et que $Z \hookrightarrow T_{n-1}$ loi de Student à $n-1$ degrés de liberté.
   - *Détermination des valeurs critiques de $Z$.*
     On cherche un réel $z_\alpha$ tel que $P(-z_\alpha \leqslant Z \leqslant z_\alpha) = 1 - \alpha$ c'est-à-dire $P(Z \leqslant z_\alpha) = 1 - \dfrac{\alpha}{2} = 0,975$.
     On obtient par lecture de la table de la loi de Student : $z_\alpha = 2,064$.
   - *Calcul de la zone d'acceptation de $H_0$.*
     $95\,\% = P(-z_\alpha \leqslant Z \leqslant z_\alpha) = P\left(\mu_0 - z_\alpha\dfrac{S}{\sqrt{n}} \leqslant \overline{X} \leqslant \mu_0 + z_\alpha\dfrac{S}{\sqrt{n}}\right)$.

     On en déduit le domaine d'acceptation :

     $$\boxed{D_0 = \left[\mu_0 - z_\alpha\frac{\widehat{\sigma}}{\sqrt{n}}\,,\ \mu_0 - z_\alpha\frac{\widehat{\sigma}}{\sqrt{n}}\right] = [58.9245\,,\ 59.0755]}$$

3. La valeur de $\overline{X}$ prise dans l'échantillon est $\overline{x_e} = 59.156$.
   On constate que $\overline{x_e} \notin D_0$. On décide de **rejeter l'hypothèse nulle $\mathbf{H}_0$** avec un risque de première espèce $\alpha = 5\%$.