Probabilités et Statistique **2011-2012**
Préorientation IMACS 2ème année

Feuille d'exercices 6
Estimation statistique

Exercice 1. Un producteur de vin vend des bouteilles de contenu déclaré de 750 ml. On suppose que la machine de remplissage automatique des bouteilles produit un contenu aléatoire (exprimé en ml) modélisé par une variable aléatoire X, de loi normale $\mathcal{N}(m, \sigma^2)$.

1. En supposant l'écart-type connu $\sigma = 20$ ml, à quelle valeur de m devrait le producteur régler sa machine si il veut être sur à 90% qu'une bouteille est remplie avec un contenu suffisant?

2. Un étudiant en statistique, en stage chez le producteur, sélectionne au hasard $n = 51$ bouteilles de vin dans la cave. Il mesure le contenu de chacune des bouteilles et trouve un contenu moyen $\bar{x}_n = 760$ ml avec un écart-type $s_n = 24.2$ ml. Construisez un intervalle de confiance pour m de niveau de confiance 90%, dans les deux cas suivants:

   (a) l'écart-type σ est supposé connu et égal à 20 ml;

   (b) l'écart-type σ est supposé inconnu.

Exercice 2. La veille des élections présidentielles on effectue un sondage auprès de n personnes sur les intentions de vote pour un candidat. On approche la proportion p de personnes dans la population totale qui ont l'intention de voter pour ce candidat, par la proportion $\hat{p}_n$ de personnes interrogées qui affirment avoir l'intention de voter pour ce candidat.

1. Déterminer le nombre minimal n de personnes que l'on doit interroger si l'on veut être sur à 95% d'estimer p avec une précision de 1%:

$$\mathbb{P}\left(|\hat{p}_n - p| \leq 1\%\right) \geq 95\%$$

   (a) en utilisant l'inégalité de Bienaymé-Tchebychev;

   (b) en utilisant le théorème central limite.

2. Un sondage effectué sur $n = 100$ personnes indique que 51% des personnes intérogées ont l'intention de voter avec ce candidat. Donner un intervalle de confiance asymptotique pour p de niveau 95%.

3. La même question si ce pourcentage est observé avec un échantillon de $n = 10000$ personnes.

Exercice 3. On suppose que la durée de vie X d'un composant électronique suit une loi exponentielle de paramètre $1/\theta$, avec $\theta > 0$ inconnu. Pour estimer θ, on mesure les durées de vie de n composants choisis au hasard. On modélise les durées de vie des n composants par des variables aléatoires $X_1, \ldots, X_n$ indépendantes et de même loi que X.

1. Donner l'espérance et la variance de X.

2. Proposer un estimateur consistant et sans biais de θ.

3. Montrer que
$$\frac{\sqrt{n}(\bar{X}_n - \theta)}{\theta} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \quad \text{lorsque } n \to \infty,$$
où $\bar{X}_n$ dénote la moyenne empirique des X_i.

4. En déduire un intervalle de confiance asymptotique pour θ de niveau 99%.

Exercice 4. Soit $X_1, \ldots, X_n$ un échantillon de variables aléatoires indépendantes de loi uniforme sur $[0, \theta]$. L'objectif est d'estimer le paramètre inconnu θ.

1. Donner l'espérance et la variance des X_i.

2. On considère l'estimateur suivant: $\hat{\theta}_n = 2\bar{X}_n$, avec $\bar{X}_n$ la moyenne empirique des X_i. Montrer que $\hat{\theta}_n$ est un estimateur consistant et non-biaisé de θ.

3. Quelle est la loi approchée de $\hat{\theta}_n$ pour n grand?

4. Construire un intervalle de confiance asymptotique pour θ de niveau 95%.