L2 Maths-Éco - Probabilités-Statistique 2 **-** 2019 – 2020

TD -Statistique - 2 bis : Estimation - suite

Exercice 1 ***Méthode des moments***
On dit que X est de loi Beta de paramètres $a, b > 0$ ($X \sim \beta(a,b)$) si X a pour densité

$$f_X(x) = C_{a,b}x^{a-1}(1-x)^{b-1}\mathbb{1}_{x\in[0,1]}.$$

1. *Montrer que $\Gamma(a)\Gamma(b)C_{a,b} = \Gamma(a+b)$ où $\forall k > 0, \Gamma(k) = \int_0^{+\infty} t^{k-1}e^{-t}\mathrm{d}t$*
2. *Exprimer $\Gamma(k+1)$ en fonction de $\Gamma(k)$.*
3. *Soit $X \sim \beta(a,b)$ et $k \geq 0$. Calculer $\mathbb{E}[X^k]$ et $\mathbb{E}[(1-X)^k]$.*
4. *En déduire $\mathbb{E}[X]$ et $Var(X)$*
5. *Estimer a, b par la méthode des moments.*

Exercice 2 ***Un raisonnement fréquentiste***
Un enseignant voudrait analyser les données suivantes : certaines erreurs impropables à des exercices (on considère une seule erreur "improbable" par exercice). On suppose qu'il y a 20 exercices dans une copie, et qu'à chaque exercice chaque étudiant a une probabilité p de faire une erreur "improbable".

1. *Quelle est la loi du nombre d'erreurs improbables par copies ? Quelles hypothèses doit-on faire pour affirmer cela ?*
2. *L'enseignant s'intéresse particulièrement aux copies 49 et 50. Quelle est la probabilité que ces deux copies aient fait toutes deux une erreur improbable au $i$eme exercice. Quelle hypothèse a-t-on fait pour affirer cela ?*
3. *Quelle est la loi du nombre d'exercices où les copies 49 et 50 ont toutes deux fait une erreur improbable (la même évidemment !) ? Quelle hypothèse fait-on pour affirmer cela ?*
4. *Au total, l'ensignant dénombres 50 erreurs "improbables" sur les 20 exercices et les 50 copies. Proposez un estimateur (en explicitant les hypothèses) $\hat{P}$ de p et un intervalle de confiance asymptotique, avec un degré de confiance de 99%. Calculer l'estimateur et l'intervalle de confiance ponctuel.*
5. *On suppose dans toute la suite que $p = \frac{1}{20}$ et on place maitenant sous l'hypothèse H_0 :"Les erreurs des différentes copies et exercices sont indépendantes". Montrer que sous H_0, la probabilité que les copies 49 et 50 aient 3 ou plus exercices avec la même erreur vaut $p_1 = \mathbb{P}\Big(Bin(20, p^2) > 2\Big)$. et calculer p_1.*
6. *L'enseignant a remarquer que les copies 49 et 50 ont fait 3 exercices avec la même erreur "improbable". Que conclure (l'enseignant ne veut pas sanctionner quelqu'un sauf s'il a un degré de confiance supérieur à $1 - \frac{1}{10^3}$?*
7. *Qu'est ce que fait l'enseignant comme grossière erreur s'il applique le raisonnement précédent en considérant toutes les paires de copies possibles, et en sanctionnant toute paire de copies ayant 3 ou plus exercices avec la même erreur "improbable" ?*
8. *L'enseignant compare maintenant les copies des personnes qui étaient assises côte à côte (50 paires au maximum). Que vaut la probabilité, sous H_0 que au moins 2 copies voisines aient au moins 3 exercices avec la même erreur ? On supposera pour cette question que la propriété "avoir 3 exercices avec une erreur en commun" est indépendante entre paires de copies. Conclure.*
9. *Critiquer les différentes hypothèses émises dans cet exercice.*

Exercice 3 ***Un raisonnment bayésien***
On se demande si Antoine et Brice ont copié, pour cela, on va comparer 2 hypothèses :

- A *: Ils ont copié*
- A^C *: Ils n'ont pas copié*

On note C *l'évènement "avoir au moins* 3 *exercices avec une erreur "improbable" en commun. On parle d'erreur "improbable" lorsque qu'il y a en moyenne environ un-e étudiant-e sur* 20 *qui la fera.*

- *Argumenter l'inégalité suivante*

$$\alpha := \mathbb{P}(C|A^C) = \mathbb{P}(Bin\big(20, \frac{1}{400} > 2\big) \approx 1.7.10^{-5}$$

- *On estime la proportion apriori de copieur potentiels à un-e étudiant-e sur* 10*. En déduire que* $p := \mathbb{P}(A) = \frac{1}{100}$*. Exprimer* $\mathbb{P}(A^C|C)$ *en fonction de* $\mathbb{P}(C|A)$*.*
- *Montrer que* $\mathbb{P}(A^C|C) \leq \frac{(1-p)\alpha}{p\mathbb{P}(C|A)} \approx \frac{1.7.10^{-3}}{\mathbb{P}(C|A)}$
- *Quelle valeur donneriez-vous à* $\mathbb{P}(C|A)$ *? Conclure en fonction du résultat*
- *Commenter et discuter les différentes hypothèses.*

Exercice 4 ***Méthode des moments***
Appliquer la méthode des moments pour estimer le ou les paramètres d'intérêt de toutes les lois suivantes :

- $Ber(p)$
- $Bin(N, p)$
- $\mathcal{G}(p)$
- $\mathcal{E}(\lambda)$
- $\mathcal{P}(\lambda)$
- $\mathcal{N}(0, 1)$