

Probabilités, Statistiques, Combinatoire : TM6

Calcul sur les lois binomiales et gaussiennes

Exercice 6.1

Simulations : Binomiales indépendantes

En réutilisant les fonctions que vous avez écrites lors des séances précédentes, écrire une fonction de simulation pour la différence de deux variables binomiales indépendantes, une de paramètres n et p , l'autre de paramètres n et q (idéalement, pour rester dans l'esprit des TM précédents : vous écrirez une fonction qui prend en paramètres l'entier n et les réels p et q , et qui retourne une fonction (sans paramètres) qui, elle, simule la loi en question).

En utilisant votre simulateur de différence de binomiales, écrire une fonction de simulation pour la variable de Bernoulli dont le paramètre (probabilité de retourner 1 plutôt que 0 est $P(n, p, q)$, défini comme la probabilité que la différence des binomiales précédentes soit (strictement) négative (on ne demande pas de *calculer* cette probabilité : seulement de simuler la loi!).

Toujours en utilisant vos fonctions issues des séances précédentes, proposer une estimation de la probabilité en question, pour les paramètres suivants : $(n, p, q) = (10, 1/5, 1/4)$; $(n, p, q) = (50, 1/5, 1/4)$; $(n, p, q) = (200, 1/5, 1/4)$.

Exercice 6.2

Calculs exacts

À l'inverse des simulations, dans cet exercice on souhaite calculer "exactement" (il y aura forcément des approximations : les calculs seront faits en utilisant le calcul sur nombres flottants) la probabilité $P(n, p, q)$.

1. Écrire une fonction `Binomiale(n,p)`, qui retourne un dictionnaire décrivant la loi binomiale de paramètres n et p (c'est-à-dire un dictionnaire dont les clés sont les valeurs possibles d'une variable aléatoire de loi binomiale de paramètres n et p , et, pour chaque clé k , la probabilité associée).
Dans la mesure du possible, évitez de passer par des calculs de factorielles ; par exemple, calculez $\mathbb{P}(B = k + 1)$ à partir de $\mathbb{P}(B = k)$.
2. Si A est binomiale de paramètres n et p , et B , indépendante de A , est binomiale de paramètres n et q , quelles sont les valeurs possibles de $C = B - A$? Écrire des fonctions permettant d'obtenir (toujours sous forme de dictionnaire) la loi de C .
3. En déduire une fonction permettant de calculer (toujours sans simulations) $P(n, p, q)$ en fonction de n , p et q . (On ne cherche pas une complexité hyper concurrentielle ; $O(n^2)$ est acceptable)
4. En déduire une fonction qui, à partir de p et q (avec $p < q$), calcule le plus petit entier n tel que $P(n, p, q)$ soit inférieur à 0.05. Comparer aux résultats théoriques obtenus en TD7 (exercice 1). (Ici, une complexité $O(n^3)$, où n est la valeur cherchée, est acceptable)

Exercice 6.3

Fonctions de répartition gaussiennes

Dans le module `math` de Python, se trouve une fonction `erf(x)` (pour “error function”) qui calcule numériquement, pour $x > 0$, la probabilité qu’une variable aléatoire gaussienne d’espérance 0 et de variance 1/2 soit comprise entre $-x$ et x ; soit, sous forme d’intégrale,

$$e(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-t^2} dt$$

(la fonction `erf` est valable aussi pour $x < 0$, et est impaire : $e(-x) = -e(x)$, la formule intégrale restant valide même quand elle ne représente pas une probabilité)

(Cet exercice est plus un exercice de TD que de TM : la difficulté est de comprendre les formules à appliquer, et pas de les programmer ; toute la difficulté de programmation se trouve dans la réalisation de la fonction `math.erf`)

1. Retrouvez (par le calcul, ou en fouillant vos notes de TD) les formules permettant d’obtenir une variable aléatoire gaussienne de paramètres m (espérance) et v (variance), à partir d’une variable normale ($m = 0$, $v = 1$), et vice-versa. Déduisez-en des formules de changement de variables entre les fonctions de répartition de ces lois gaussiennes (on rappelle que la fonction de répartition d’une variable X , est définie par $F_X(x) = \mathbb{P}(X \leq x)$; pour une variable gaussienne, il s’agit d’une fonction continue sur \mathbb{R}).
2. À partir de la fonction `math.erf`, écrivez une fonction `Fnormale(x)` qui, appelée avec un réel x , retourne la probabilité qu’une variable aléatoire normale (donc gaussienne d’espérance 0 et de variance 1) soit inférieure ou égale à x .
Pour cela, commencez par vérifier que $(1+\text{math.erf}(x))/2$ représente la fonction de répartition d’une gaussienne d’espérance 0 et de variance 1/2 (utilisez la propriété de symétrie des gaussiennes) ; et, en exprimant une formule permettant d’obtenir une gaussienne d’espérance 0 et de variance 1/2 à partir d’une variable normale, exprimez la fonction de répartition de la variable normale en fonction de la fonction d’erreur.
3. À partir de votre fonction `Fnormale`, écrivez une fonction `Fgaussienne(m,v,x)`, qui retourne la valeur en x de la fonction de répartition d’une gaussienne d’espérance m et de variance $v > 0$. (Il s’agit ici encore de faire les bons changements de variable à partir de la fonction `Fnormale` ; il n’est pas nécessaire d’écrire ni de calculer d’intégrales)
4. Enfin, écrivez une fonction qui, prenant en entrée deux réels m et v (avec $v > 0$, et deux autres réels a et b (avec $a \leq b$), retourne la probabilité qu’une gaussienne d’espérance m et de variance v , prenne une valeur entre a et b .