Probabilité et statistiques

L2 prépa

William Hodonou

Novembre 2020

Outline

Introduction

En réalité, on se rend compte que la grande majorité des phénomènes statistiques peuvent être décrits par un nombre réduit de modèles probabilistes. Il importe dans un premier temps de pouvoir décrire de façon adéquate le mécanisme du processus réel étudié (temps d'attente, nombre de passages dans un intervalle de temps, nombre d'essais avant d'obtenir tel résultat, etc.). Dans un second temps, une fois cette caractérisation réalisée, nous pouvons choisir la loi théorique qui paraît le mieux convenir pour modéliser le phénomène observé, l'étape suivante consistant à estimer les paramètres de la loi.

Il est donc nécessaire de connaître les principales lois de probabilité, tant du point de vue de leur expression, de leurs caractéristiques de position et de dispersion, que du type de phénomène qu'elles prétendent modéliser. On va différencier ici les lois selon leur nature discrète ou continue.

Outline

Loi Binomiale

La loi binomiale décrit le nombre d'apparitions d'un évènement (disons "succès") dans une suite de n épreuves indépendantes. La probabilité d'apparition de l'évènement à une épreuve est p (a contrario la probabilité de non-apparition est $q = (1-p)$, elle est constante tout le long de l'expérience. la probabilité s'écrit: $X \sim B(n,p)$

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \ \ \forall\ x = 0, 1, 2, ..., n$$

$\binom{n}{x} = \frac{n!}{x!(n-x)!}$

La loi binomiale $X \sim B(n,p)$ est définie par le paramètre p, qui indique la probabilité d'apparition de l'évènement pour une épreuve ; n est le nombre d'épreuves réalisées, c'est une donnée du problème, nous n'avons pas à l'estimer.

Loi Binomiale

Caractéristiques de la loi

$$E(X) = np$$

$$V(X) = np(1-p)$$

$$M_x(t) = (1 - p + pe^t)^n$$

Loi Binomiale

Exemple:

Un étudiant passe un examen à choix multiples avec n= 10 questions. Il n'a pas assisté à la classe ou étudié pendant trois semaines et prévoit de deviner sur chaque question en demandant à sa calculatrice d'afficher un entier aléatoire de 1 à 5. (Il y a 5 choix pour chaque question.) Soit X le nombre de questions sur les 10 pour lesquels l'étudiant devine correctement. Déterminer la probabilité que l'étudiant qui a deviné sur toutes les 10 questions n'obtienne que 2 bonnes réponses.

Loi de Poisson

Cette loi est très utilisée car elle apparaît dans les processus de comptage d'évènements pouvant survenir pendant une unité de temps donnée (sous certaines hypothèses). On se sert donc de cette loi lorsqu'on s'intèresse au nombre de personnes sortant d'une salle d'attente, arrivant à un guichet, passant à un péage, au nombre de coups de fil recus, etc. C'est également la seule loi discrète à avoir son espérance égale à sa variance.on note $X \sim p(\lambda)$

$$P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}$$

avec $\lambda \geq 0$

Loi de Poisson

Caractéristiques de la loi

$$E(X) = \lambda = V(X)$$

$$M_x(t) = e^{\lambda(e^{t-1})}$$

Loi de Poisson

Exemple:

Les titulaires d'une police d'assurance déposent des réclamations à un taux moyen de 0,45 par an. Utilisez le modèle de Poisson pour répondre aux questions suivantes. (a) Trouvez la probabilité qu'un preneur d'assurance dépose au moins une réclamation par an. (b) Trouvez le nombre moyen de réclamations par preneur d'assurance et par an. (c) Supposons que chaque réclamation paie exactement 1 000 \$. Trouvez le montant moyen de la réclamation pour un preneur d'assurance en un an.

Loi Géométrique

On exécute une série d'épreuves indépendantes ayant chacune la probabilité p d'être un succès, $0 < p < 1$, jusqu'à obtenir le premier succès. Si l'on désigne le nombre d'épreuves nécessaires jusqu'à ce résultat par X on aura:

$$p(X = x) = p(1 - p)^{n-1} \ \ \forall \ x = 1, 2, 3, ..., n$$

En effet, pour que X prenne n pour valeur, il faut et suffit que les n — 1 premières épreuves soient des échecs tandis que la n-ième devra être un succès.

Loi Géométrique

Caractéristiques de la loi

$$E(X) = \frac{1}{p}$$

$$V(X) = \frac{1-p}{p^2}$$

$$M_x(t) = \frac{p}{1-(1-p)e^t}$$

Loi Géométrique

Exemple:

Si la probabilité est de 0,75 qu'un candidat à un permis de conduire réussisse l'examen sur route à un essai donné, quelle est la probabilité qu'un candidat réussisse finalement l'examen au quatrième essai?

Loi Hypergéométrique

Pour préciser les idées, nous dirons que nous sommes en présence d'une urne contenant M boules, K sont de couleur blanche qui nous intéresse, M-K sont de couleur noire. La loi hypergéométrique décrit le nombre X de boules blanches extraites de l'urne si nous effectuons un tirage exhaustif (sans remise) de n boules.

$$p(x) = \frac{\binom{K}{X}\binom{M-K}{n-x}}{\binom{M}{n}}$$

Loi Hypergéométrique

Caractéristiques de la loi

$$E(X) = \frac{nK}{M}$$

$$V(X) = \frac{nK(M-K)(M-n)}{M^2(M-1)}$$

Loi Hypergéométrique

Exemple:

Sur un groupe de 25 ouvriers d'usine, 20 sont à faible risque et cinq à haut risque. Deux des 25 travailleurs de l'usine sont choisis au hasard sans remplacement.
a- Calculez la probabilité qu'un seul des deux ouvriers d'usine sélectionnés présente un risque faible.

Outline

Lois Continues

Les lois continues sont caractérisées à l'aide des fonctions de densité. Théoriquement, une v.a. continue doit pouvoir prendre n'importe quelle valeur dans un intervalle donné. C'est pour cela d'ailleurs que la probabilité ponctuelle est par définition nulle. En réalité, dans certains cas, la distinction loi discrète - loi continue est essentiellement formelle. Il est possible d'approximer certaines lois discrètes à l'aide de lois continues.

Loi Uniforme continue

La v.a. X suit une loi uniforme sur $[a, b]$ si sa densité de probabilité est donnée par :

$$f(x) = \begin{cases} \frac{1}{b-a}, & si\ a \leq x \leq b \\ 0, & sinon \end{cases}$$

Dans le cas de la loi uniforme, il est aisé de calculer la fonction de répartition :

$$F(x) = \frac{x-a}{b-a}$$

Loi Uniforme continue

Caractéristiques de la loi

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

$$M_x(t) = \frac{e^{bt} - e^{at}}{(b-a) \times t}$$

Loi Uniforme continue

Exemple:

Une entreprise s'attend à recevoir le paiement d'une facture importante aujourd'hui. Le temps X jusqu'à ce que le paiement soit reçu est uniformément réparti sur l'intervalle $[1, 9]$, entre 1 et 9 heures à partir de maintenant, tous les instants de l'intervalle étant également probables.Donner la fonction de densité pour X. Donner la probabilité que l'heure de réception soit comprise entre 2 et 5 heures à partir de maintenant. Donner la fonction de répartition.

Loi exponentielle

Cette loi est très souvent utilisée en fiabilité par exemple. En fait elle modélise tout ce qui est **durée de vie** : d'une ampoule, d'un système informatique, d'un noyau atomique, d'un impact d'une annonce d'une nouvelle mesure économique, etc. On peut également la caractériser par le fait qu'il s'agit de l'unique loi sans mémoire continue. Cela signifie que la probabilité de b vivre encore T période sachant qu'on a déjà vécu T' est rigoureusement identique avec la probabilité de vivre T. La fonction de densité est

$$f(x) = \lambda e^{-\lambda x}$$

λ est le seul paramètre de la loi. La fonction de répartition est facilement obtenu :

$$F(x) = 1 - e^{-\lambda x}$$

Loi exponentielle

Caractéristiques de la loi

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

pour tout $t < \lambda$

Loi normale

- La loi normale est aussi appelée loi de Laplace-Gauss, ou loi gaussienne.
- Loi essentielle pour la statistique inférentielle.

Définition

X suit une loi normale de paramètres μ et σ si sa densité de probabilité est donnée par :

$$(\forall x \in \mathbb{R}) \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

on note $X \sim N(\mu, \sigma)$

Loi normale

Caractéristiques de la loi

$$E(X) = \mu$$

$$V(X) = \sigma^2$$

Loi normale

Loi normale centrée réduite

Il existe une v.a. dite centrée et réduite U, qui suit une loi normale qui est associée à chaque valeur de X. Pour passer de X à U, on effectue un changement de variable en posant :

$$U = \frac{X - \mu}{\sigma}$$

Dans ce cas, la densité de probabilité de U est

$$f(U) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{U}{2}}$$

Pour passer de X à U, on peut dire que tout se passe comme si $\mu = 0$ et $\sigma = 1$. Les valeurs de F(U) sont tabulées. F(U) est la fonction de répartition de U. Si on connait U, on peut calculer F(U) Si on connait F(U), on peut calculer U