

Statistiques appliquées à la gestion

Licence 2 gestion
2023/2024

R. ZAFRI

Maître de Conférences

raouf.zafri@univ-paris1.fr

Sommaire du cours

Chapitre 1 : Dénombrements et analyse combinatoire

Chapitre 2 : Introduction aux formalismes des probabilités

Chapitre 3 : Indépendances et probabilités conditionnelles

Chapitre 4 : Introduction aux variables aléatoires réelles

Chapitre 5 : Variables aléatoires discrètes usuelles

Chapitre 6 : Variables aléatoires réelles absolument continues

Chapitre 7 : Loi de probabilité

Chapitre 5. Variables aléatoires discrètes usuelles

8. La loi hypergéométrique

9. Convergence en loi

8. La loi hypergéométrique

A. Définition

La loi hypergéométrique concerne les situations où on sélectionne des éléments **sans remise** dans un ensemble fini. Cela signifie que chaque fois que l'on effectue une sélection, on réduit la taille de l'ensemble pour les sélections suivantes. Ainsi, la probabilité de sélection change à chaque tirage.

8. La loi hypergéométrique

A. Définition

Exemple

Nous avons une boîte de bonbons contenant :

8 bonbons au chocolat et 12 bonbons à la fraise (objet)

Au total, on a donc 20 bonbons (population totale).

On décide de fermer les yeux et de prendre au hasard 5 bonbons de la boîte (échantillon), tout cela en une seule fois et sans remise. Maintenant, on se demande : "Quelle est la probabilité que parmi ces 5 bonbons, exactement 3 soient des bonbons au chocolat ?«

La variable aléatoire X : nombre de bonbons au chocolat obtenus parmi les 5 tirés

La loi hypergéométrique permet de calculer cette probabilité en tenant compte du nombre total de bonbons dans la boîte, du nombre de chaque type de bonbon, et du nombre de bonbons qu'on sélectionne.

- L'objet est partitionné en un évènement et son contraire (succès et échec) (chocolat, non chocolat)
- L'expérience est menée 5 fois : $n = 5$ (on interdit la répétition, les issues de l'expérience sont des combinaisons)
- Au bout des 5 essais, la variable aléatoire comptabilise le nombre de succès obtenus

La loi de X est donc $H(20 ; 5 ; \frac{8}{20})$

8. La loi hypergéométrique

B. La loi de probabilité de la loi hypergéométrique

Une variable aléatoire X discrète suit une loi hypergéométrique de paramètres $H(N, n, p)$ si sa loi de probabilité est modélisée par la fonction suivante :

$$P(X = k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}$$

- N : la taille totale de la population
- Np : le nombre d'éléments favorables dans la population
- Nq : = le nombre d'éléments non favorables dans la population
- n : est la taille de l'échantillon que nous sélectionnons
- k est le nombre d'éléments favorables que nous voulons sélectionner dans l'échantillon

8. La loi hypergéométrique

Exemple

Une urne contenant 10 billes, parmi lesquelles 4 sont rouges et 6 sont bleues. Vous voulez en choisir 3 au hasard, sans remise.

Quelle est la probabilité que vous choisissiez exactement 2 billes rouges dans vos 3 sélections ?

Résolution

Dans cet exemple, la taille de la population N est de 10 (le nombre total de billes dans l'urne), le nombre d'éléments favorables est de 4 (le nombre de billes rouges) donc $p = 4/10$, la taille de l'échantillon n est de 3 (le nombre de billes que vous choisissez), et nous voulons connaître la probabilité d'obtenir exactement $k=2$ billes rouges dans notre sélection.

Ce problème peut être modélisé par une loi hypergéométrique de paramètres $H(10, 3, \frac{4}{10})$.

Utilisons la formule de la loi hypergéométrique :
$$P(X = 2) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}} = \frac{\binom{4}{2} \binom{6}{1}}{\binom{10}{3}} = 0,3$$

Donc, la probabilité d'obtenir exactement 2 billes rouges dans votre sélection de 3 billes est de 0,3, soit 30%.

8. La loi hypergéométrique

C. Les caractéristiques d'une variable hypergéométrique

Soit la variable aléatoire X qui suit une loi hypergéométrique de paramètres $H(N, n, p)$

Si $X \rightarrow H(N, n, p)$ alors

$$E(X) = np$$

$$V(X) = npq \frac{N-n}{N-1}$$

8. La loi hypergéométrique

Exemple

Une entreprise de fabrication produit des lots de 1000 pièces d'équipement. On estime qu'en moyenne, 5% des pièces d'un lot peuvent présenter un défaut de fabrication. L'équipe de contrôle qualité effectue un échantillonnage aléatoire sans remise de 60 pièces pour inspection.

1. Quelle est la probabilité que l'échantillon contienne exactement 3 pièces défectueuses ?
2. Quelle est l'espérance du nombre de pièces défectueuses trouvées dans l'échantillon ?
3. Calculer la variance du nombre de pièces défectueuses dans l'échantillon.

Résolution

- Pour calculer cette probabilité, nous utilisons la loi hypergéométrique. Soit X le nombre de pièces défectueuses dans l'échantillon :
- $N=1000$ (taille du lot),
- $p = 0,05$ (proportion attendue de pièces défectueuses)
- $q= 0,95$
- $n=60$ (taille de l'échantillon),
- $k=3$ (nombre de pièces défectueuses à trouver dans l'échantillon).

8. La loi hypergéométrique

Résolution

1) Ce problème peut être modélisé par une loi hypergéométrique de paramètres $H(1000, 60, 0.5)$

$$P(X = k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}$$
$$P(X = 3) = \frac{\binom{50}{3} \binom{950}{57}}{\binom{1000}{60}} \approx 0,237$$

$$Np = 0,05 \times 1000 = 50$$

$$Nq = 1000 - 50 = 950$$

2) L'espérance $E(X)$ pour une distribution hypergéométrique est calculée avec la formule : $E(X) = np = 60 \times 0,05 = 3$. On peut s'attendre donc à trouver 3 pièces défectueuses dans l'échantillon.

3) La variance $V(X)$ pour une distribution hypergéométrique est donnée par : $V(X) = npq \frac{N-n}{N-1} = 60 \times 0,05 \times 0,95 \times \frac{1000-60}{1000-1} \approx 2.682$

La variance est d'environ 2.682, ce qui nous informe sur la dispersion du nombre de pièces défectueuses trouvées dans l'échantillon par rapport à la moyenne calculée.

8. La loi hypergéométrique

D. Exemples concrets d'utilisation de la loi géométrique en management

- Gestion des ressources humaines et recrutement : La loi hypergéométrique peut être utilisée pour calculer les probabilités de sélectionner un certain nombre de candidats ayant les compétences requises dans un groupe de candidats. Cela est particulièrement utile dans les processus de recrutement où les entreprises cherchent à maximiser la probabilité de sélectionner les candidats les plus qualifiés parmi un grand nombre de postulants.
- Contrôle de qualité : Dans le cadre du contrôle de qualité, cette loi permet d'évaluer les risques d'erreur dans les lots de production. Par exemple, si une entreprise veut s'assurer que la proportion de produits défectueux dans un lot ne dépasse pas un certain seuil, la loi hypergéométrique peut aider à déterminer la taille de l'échantillon nécessaire pour tester cette hypothèse avec un niveau de confiance donné.
- Gestion des stocks et de la chaîne d'approvisionnement : La loi hypergéométrique peut être utilisée pour évaluer les probabilités associées à la présence ou à l'absence de produits spécifiques dans un inventaire, ce qui est crucial pour la gestion des stocks et la prévention des ruptures de stock. Elle aide à prendre des décisions éclairées sur la quantité de stock à maintenir pour répondre à la demande sans encourir des coûts de stockage excessifs.
- Gestion de projet et planification stratégique : Dans la gestion de projet, cette loi peut aider à évaluer les risques associés à la répartition des ressources. Par exemple, elle peut être utilisée pour déterminer la probabilité que suffisamment de membres d'une équipe aient les compétences nécessaires pour mener à bien un projet dans les délais impartis, en tenant compte des contraintes de ressources.

8. La loi hypergéométrique

Exercice

L'oral d'un concours comporte au total 100 sujets; les candidats tirent au sort 3 sujets et choisissent alors le sujet traité parmi ces trois sujets. Un candidat se présente en ayant révisé 60 sujets sur les 100.

1. Quelle est la probabilité pour que le candidat ait révisé:
 - a) les trois sujets tirés ;
 - b) exactement deux sujets sur les trois sujets;
 - c) aucun des trois sujets.
2. Définir une variable aléatoire associée à ce problème et donner sa loi de probabilité, son espérance.

8. La loi hypergéométrique

Corrigé

1. La variable aléatoire associée à ce problème est X «nombre de sujets révisés parmi les 3» ; son support est l'ensemble $\{0,1,2,3\}$. La loi de X est une loi hypergéométrique puisque l'événement $[X = k]$, pour k compris entre 0 et 3, se produit si le candidat tire k sujet(s) parmi les 60 révisés, et $3-k$ sujets parmi les 40 non révisés. Alors :

a) Les trois sujets tirés ont été révisés : $P(X = 3) = \frac{\binom{60}{3}\binom{40}{0}}{\binom{100}{3}}$

b) Deux des trois sujets tirés ont été révisés : $P(X = 2) = \frac{\binom{60}{2}\binom{40}{1}}{\binom{100}{3}}$

c) Aucun des trois sujets : $P(X = 0) = \frac{\binom{60}{0}\binom{40}{3}}{\binom{100}{3}} = 6,110 \times 10^{-2}$

8. La loi hypergéométrique

Corrigé

2. La loi de probabilité de X est donnée sur le support $\{0,1,2,3\}$ par : $P(X = k) = \frac{\binom{60}{k} \binom{40}{3-k}}{\binom{100}{3}}$

$$k = 0 : P(X = 0) \approx 6.110 \times 10^{-2}$$

$$k = 1 : P(X = 1) \approx 0.289$$

$$k = 2 : P(X = 2) \approx 0.438$$

$$k = 3 : P(X = 3) \approx 0.212$$

L'espérance est $E(X) = 1.8$ (selon la formule $E(X) = np$).

8. La loi hypergéométrique

Exercice

On prend au hasard, en même temps, 3 ampoules dans un lot de 15 dont 5 sont défectueuses. Calculer la probabilité des événements:

- A. au moins une ampoule est défectueuse;
- B. les 3 ampoules sont défectueuses;
- C. exactement une ampoule est défectueuse

8. La loi hypergéométrique

Corrigé

On utilise une loi hypergéométrique

$$P(A) = 1 - \frac{\binom{5}{0}\binom{10}{3}}{\binom{15}{3}} = 0.736$$

$$P(B) = \frac{\binom{5}{3}\binom{10}{0}}{\binom{15}{3}} = 2.197 \times 10^{-2}$$

$$P(C) = \frac{\binom{5}{1}\binom{10}{2}}{\binom{15}{3}} = 0.494$$

9. Convergence en loi

A. Approximation d'une loi binomiale par la loi Poisson

- Soit X une variable aléatoire suivant la loi binomiale $B(120; 0,08)$. On souhaite calculer différentes valeurs de $P(X = k)$ pour k compris entre 0 et 120.
- On sait que $P(X = k) = \binom{120}{k} 0,08^k (1 - 0,08)^{120-k}$
- Le problème est qu'un tel calcul peut s'avérer difficile à obtenir avec une calculatrice « traditionnelle ». Par exemple, dans le cas où $k = 60$, il faudrait déterminer $\binom{120}{60} = 96614908840363322603893139521372656$ s'en suivrait le calcul de $0,08^{60}$... Nous vient alors l'idée d'une approximation qui rendrait le calcul plus « simple ».
- Commençons par remarquer que la probabilité de succès est très faible (0,08). Le succès peut donc être considéré comme un « événement rare »
- On pense alors à une approximation par une loi de Poisson (loi des événements rares).
- $E(X) = 120 \times 0,08 = 9,6$ et, d'après ce qui précède, on sait qu'une loi de Poisson de paramètre 9,6 a cette espérance.
- Comparons donc les distributions des lois $B(120 ; 0,08)$ et $P(9,6)$

9. Convergence en loi

A. Approximation d'une loi binomiale par la loi Poisson

Soit X une variable aléatoire suivant la loi binomiale $B(n ; p)$. On suppose que :

- $n \geq 30$
- $p \leq 0,1$
- $np < 10$

On considère la variable aléatoire $Y \rightarrow P(np)$

Alors, pour tout entier naturel k compris entre 0 et n , on a : $P(X = k) \approx P(Y = k)$

9. Convergence en loi

Application

On considère un jeu auquel la probabilité de gagner une partie est égale à 3%. On effectue 200 parties de manière indépendante.

On s'intéresse à la probabilité de gagner exactement 10 parties. Soit X la variable aléatoire donnant le nombre de parties gagnées sur les 200. On sait qu'alors :

$X \rightarrow B(200; 0,03)$. On a $E(X) = 200 \times 0,03 = 6$.

Les hypothèses du théorème sont vérifiées puisque :

$n = 200 \geq 30, p = 0,03 \leq 0,1$ et $np = 6 < 10$. On considère alors la variable aléatoire $Y \rightarrow P(6)$.

Ainsi : $P(X = 10) \approx P(Y = 10) \approx 0,041$ (par lecture de la table de la loi de Poisson). Ainsi, la probabilité de gagner exactement 10 fois est environ égale à 4,1%

9. Convergence en loi

Table de la loi de Poisson

$k \backslash \lambda$	1	1.5	2	3	4	5	6	7	8	9	10
0	0.368	0.223	0.135	0.050	0.018	0.007	0.002	0.001	0.000	0.000	0.000
1	0.368	0.335	0.271	0.149	0.073	0.034	0.015	0.006	0.003	0.001	0.000
2	0.184	0.251	0.271	0.224	0.147	0.084	0.045	0.022	0.011	0.005	0.002
3	0.061	0.126	0.180	0.224	0.195	0.140	0.089	0.052	0.029	0.015	0.008
4	0.015	0.047	0.090	0.168	0.195	0.176	0.134	0.091	0.057	0.034	0.019
5	0.003	0.014	0.036	0.101	0.156	0.176	0.161	0.128	0.092	0.061	0.038
6	0.001	0.004	0.012	0.050	0.104	0.146	0.161	0.149	0.122	0.091	0.063
7	0.000	0.001	0.003	0.022	0.060	0.104	0.138	0.149	0.140	0.117	0.090
8		0.000	0.001	0.008	0.030	0.065	0.103	0.130	0.140	0.132	0.113
9			0.000	0.003	0.013	0.036	0.069	0.101	0.124	0.132	0.125
10				0.001	0.005	0.018	0.041	0.071	0.099	0.119	0.125
11				0.000	0.002	0.008	0.023	0.045	0.072	0.097	0.114
12					0.001	0.003	0.011	0.026	0.048	0.073	0.095
13					0.000	0.001	0.005	0.014	0.030	0.050	0.073
14						0.000	0.002	0.007	0.017	0.032	0.052
15							0.001	0.003	0.009	0.019	0.035
16							0.000	0.001	0.005	0.011	0.022
17								0.001	0.002	0.006	0.013
18								0.000	0.001	0.003	0.007
19									0.000	0.001	0.004
20										0.001	0.002
21										0.000	0.001
22											0.000

9. Convergence en loi

Exercice

Un agent immobilier a estimé que la probabilité de vendre un appartement suite à une visite était de 8%. Il effectue en général 100 visites par mois.

On considère que les visites d'appartement sont des expériences aléatoires indépendantes les unes des autres.

On appelle A la variable aléatoire égale au nombre d'appartements vendus en un mois après une visite.

1. Justifier que la variable aléatoire A suit une loi binomiale dont on précisera les paramètres.
2. On souhaite calculer la probabilité que l'agent vende 10 appartements en un mois après une visite.
 - Déterminer $\binom{100}{10}$ à l'aide de la calculatrice. Que peut-on en conclure ?
 - Justifier que l'on peut approcher la loi de X par une loi de Poisson que l'on précisera
 - Grâce à cette approximation, déterminer la probabilité voulue

9. Convergence en loi

Corrigé

1. On a affaire à une expérience aléatoire admettant exactement deux issues :

succès : « vendre un appartement suite à une visite » de probabilité 0,08

échec : « ne pas vendre un appartement suite à une visite » de probabilité 0,92

- Cette expérience aléatoire est donc une épreuve de Bernoulli. On répète cette expérience aléatoire 100 fois de manière indépendante : il s'agit donc d'un schéma de Bernoulli.
- La variable aléatoire X est égale au nombre d'appartements vendus après une visite en un mois, c'est-à-dire au nombre de succès obtenus à l'issue des 100 épreuves de Bernoulli. Par conséquent, $X \rightarrow B(100; 0,08)$.

9. Convergence en loi

2. À l'aide de la calculatrice, on obtient $\binom{100}{10} \approx 1,731030946 \times 10^{13}$

Ce nombre est difficile à obtenir car il fait intervenir des factorielles importantes (il vaut en réalité 17310309456440). La calculatrice n'en donne donc qu'une valeur approchée. Il en serait de même pour $0,08^{10}$ et $0,92^{90}$ qui sont les deux autres facteurs intervenants dans le calcul de $P(X = 10)$. Le recours à une approximation de la loi binomiale semble alors nécessaire.

- Comme :
- $n = 100 \geq 30$
- $p = 0,08 \leq 0,1$
- $np = 100 \times 0,08 = 8 < 10$

on peut approcher la loi de A par la loi $P(8)$

- Par lecture de la table (colonne $\lambda = 8$ et ligne $k = 10$), on obtient : $P(A = 10) \approx 0,099$ (une valeur approchée à 10^{-3} près de la probabilité obtenue en utilisant la loi binomiale est 0,1024 soit une erreur commise d'environ 3%)

9. Convergence en loi

B. Approximation de la loi hypergéométrique par la loi binomiale

La loi hypergéométrique (loi d'une variable aléatoire lors d'un tirage sans remise) peut être approchée par la loi binomiale lorsque le nombre d'individus de la population est très grand devant le nombre d'individus étudiés.

Exemple

Lorsque le nombre de tirages n est très petit par rapport au nombre N de boules dans l'urne. Alors, à chaque tirage, la proportion de boules blanches est pratiquement inchangée. Donc ceci revient pratiquement à effectuer des tirages avec remise. En pratique, on utilise l'approximation $P(H(N, n, p) = k) \simeq P(B(n, p) = k)$ dès que $\frac{n}{N} \leq 0,1$.

9. Convergence en loi

Exercice

Vous gérez une chaîne de magasins de vêtements avec un stock centralisé. Dans ce stock, vous avez 200 chemises, dont 40 sont de marque premium. Pour une vente spéciale, vous décidez de sélectionner aléatoirement sans remise 10 chemises. Vous souhaitez connaître la probabilité de sélectionner exactement 3 chemises premium dans cet échantillon de 10 chemises.

1. On appelle X la variable aléatoire égale au nombre de chemises premium.
Calculer la probabilité d'obtenir exactement 3 chemises premium premium.
2. Justifier que l'on peut approcher la loi de A par une loi binomiale que l'on précisera

9. Convergence en loi

Exercice

$$X \rightarrow H(200, 10, 0,2)$$
$$P(X = k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}$$
$$P(X = 3) = \frac{\binom{40}{3} \binom{160}{7}}{\binom{200}{10}} \approx 20,52\%$$

$$p = \frac{40}{200} = 0,2$$
$$Np = 40$$
$$Nq = 200 - 40 = 160$$

L'approximation binomiale est considérée comme adéquate si la taille de l'échantillon n est petite par rapport à la taille de la population N , et si la proportion de succès $p=K/N$ reste constante. Ici : $p=40/200 = 0.2$ et $n/N = 10/200 = 0,05 < 0,1$.

On peut utiliser l'approximation $P(H(N, n, p) = k) \simeq P(B(n, p) = k)$

Dans notre cas : $P(H(200, 10, 0,2) = k) \simeq P(B(10, 0,2) = k)$

$$P(X = k) = \binom{10}{k} (0,2)^k (1 - 0,2)^{10-k}$$

$$P(X = 3) = \binom{10}{3} (0,2)^3 (1 - 0,2)^{10-3} = 21,50\%$$

9. Convergence en loi

Exercice

Dans une population très nombreuse, on rencontre en moyenne 0,4% de non voyants.

- 1) Dans un échantillon de 100 personnes, quelle est la probabilité de n'avoir aucun non voyant ?
- 2) Peut-on utiliser la loi Poisson pour répondre à cette question

9. Convergence en loi

Résolution

Dans une population très nombreuse, on rencontre en moyenne 0,4% de non voyants.

1) Dans un échantillon de 100 personnes, quelle est la probabilité de n'avoir aucun non voyant ?

On répète 100 fois l'expérience : choisir un individu au hasard avec une probabilité $p = 0,4\% = 0,004$

Qu'il soit non voyant (succès). La population étant très grande, on considère que p est invariable, ce qui fait que le tirage est assimilé « avec remise » et qui nous autorise une loi binomiale.

Soit X la variable aléatoire qui compte le nombre de non voyants sur les 100 personnes. La loi de X est $B(100 ; 0,004)$.

$$P(X = 0) = \binom{100}{0} (0,004)^0 (1 - 0,004)^{100-0} = 0,6698$$

9. Convergence en loi

Résolution

2) Peut-on utiliser la loi de Poisson pour répondre à cette question

- $n = 100 \geq 30$
- $p = 0,004 \leq 0,1$
- $np = 100 \times 0,004 = 0,4 < 10$

Donc $B(100 ; 0,004)$ est à peu près $P(0,4)$

D'après la loi de Poisson ($\lambda = 0,4$), $P(X = 0) \approx 0,670$

Synthèses des lois usuelles discrètes

Nom	Paramètres	Notation	Valeurs	Loi	$E(X)$	$V(X)$
Uniforme	$n \in \mathbb{N}^*$	$\mathcal{U}(n)$	$\llbracket 1, n \rrbracket$	$P(X = k) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Bernoulli	$p \in]0, 1[$	$\mathcal{B}(1, p)$	$\{0, 1\}$	$P(X = 1) = p$	p	pq
Binomiale	(n, p)	$\mathcal{B}(n, p)$	$\llbracket 0, n \rrbracket$	$P(X = k) = \binom{n}{k} p^k q^{n-k}$	np	npq
Géométrique	p	$\mathcal{G}(p)$	$\mathbb{N}^*$	$P(X = k) = pq^{k-1}$	$\frac{1}{p}$	$\frac{q}{p^2}$
Hypergéométrique	(N, n, p)	$\mathcal{H}(N, n, p)$	$\subset \llbracket 0, n \rrbracket$	$P(X = k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}$	np	$npq \frac{N-n}{N-1}$
Poisson	$\lambda \in \mathbb{R}_+$	$\mathcal{P}(\lambda)$	$\mathbb{N}$	$P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	λ	λ